REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 16	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) TOPICS IN TIME DELAY ESTIMATION		5. TYPE OF REPORT & PERIOD COVERED TECHNICAL REPORT Sept. 1982 – Oct. 1984
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Alfred O. Hero & Stuart C. Schwartz		8. CONTRACT OR GRANT NUMBER(s) N00014-81-K-0146
9. PERFORMING ORGANIZATION NAME AND ADDRESS Information Sciences & Systems Laboratory Dept. of Electrical Eng. & Computer Sci. Princeton University, Princeton, NJ 08544		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS NR SRO-103
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research (Code 411SP) Department of the Navy Arlington, Virginia 22217		12. REPORT DATE February 1985
		13. NUMBER OF PAGES 215
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

   Approved for public release; distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

   Also submitted as the Ph.D. Thesis of Alfred Hero to the EECS Department,
   Princeton University, Princeton, NJ 08544, October 1984.

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

   time delay estimation
   robust Wiener filter
   center of symmetry estimator
   level-crossing
   inhomogeneous Poisson

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

      Time delay estimation forms the base from which passive
   localization and tracking may function accurately.   In this
   thesis we consider both the practical and theoretical aspects of
   time delay estimation when classical optimal estimation tech-
   niques break down. Of special concern are two adverse situations:
   the underlying statistics of the observations are incompletely
   specified; and the observations contain a narrowband component. (Continued)

DD FORM 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-LF-014-6601

Abstract Continued

For the former we take a min-max approach: the signal and noise spectra are assumed to belong to uncertainty classes of spectra and a processor is specified which attains the best possible performance over the entire class. While the classical min-max time delay estimation problem is solved only in a restricted sense, use is made of the existing theory on Robust (min-max) Wiener filtering to motivate an approximate min-max processor. Performance curves for the uncertain spectra situation show that the approximate scheme has the potential to outperform the Generalized Cross Correlator (GCC) class of optimal processors.

It is well known that as the observations become narrowband the classical optimal time delay estimate degenerates rapidly due to the so called "peak ambiguity" problem. For Gaussian observations in the narrowband case, this arises from the multipeaked nature of the likelihood function over which the global maximum is acquired. We are interested in two aspects of the problem: the improvement of performance, where possible, through alternative estimation techniques; and the characterization of the large error performance of the GCC class. Two new processors are presented, one arising from constraining the optimal estimator to have low sidelobes, and the other derived from an optimization over a new family of processors, the Center of Symmetry Estimator (CSE). Simulation results for the CSE indicate that it can have significantly better large error performance than the GCC. We approach the performance characterization problem in the following way: the performance of the optimal estimator in the presence of large errors is related to the probability that the likelihood function gives an estimate which is far removed from the true time delay. This probability, in turn, admits a representation in terms of level up-crossing probabilities. In Chapter 4, we present a general theorem on the asymptotic behavior of the level up-crossings of a non-stationary random process. Then, based on the results of Chapter 4, a Poisson model is invoked as an approximation to the level crossings and a numerical study of the large error probability and the global variance is undertaken for lowpass and bandpass signal spectra.

TOPICS IN TIME DELAY ESTIMATION

A.O. HERO and S.C. SCHWARTZ

TOPICS IN TIME DELAY ESTIMATION

A.O. Hero and S.C. Schwartz
Department of Electrical Engineering
and Computer Science
Princeton University
Princeton, New Jersey 08544

Abstract

Time delay estimation forms the base from which passive localization and tracking may function accurately. In this thesis we consider both the practical and theoretical aspects of time delay estimation when classical optimal estimation techniques break down. Of special concern are two adverse situations: the underlying statistics of the observations are incompletely specified; and the observations contain a narrowband component.

For the former we take a min-max approach: the signal and noise spectra are assumed to belong to uncertainty classes of spectra and a processor is specified which attains the best possible performance over the entire class. While the classical min-max time delay estimation problem is solved only in a restricted sense, use is made of the existing theory on Robust (min-max) Wiener filtering to motivate an approximate min-max processor. Performance curves for the uncertain spectra situation show that the approximate scheme has the potential to outperform the Generalized Cross Correlator (GCC) class of optimal processors.

It is well known that as the observations become narrowband the classical optimal time delay estimate degenerates rapidly due to the so called "peak ambiguity" problem. For Gaussian observations in the narrowband case, this arises from the multipeaked nature of the likelihood function over which the global maximum is acquired. We are interested in two aspects of the problem: the improvement of performance, where possible, through alternative estimation techniques; and the characterization of the large error performance of the GCC class. Two new processors are presented, one arising from constraining the optimal estimator to have low sidelobes, and the other derived from an optimization over a new family of processors, the Center of Symmetry Estimator (CSE). Simulation results for the CSE indicate that it can have significantly better large error performance than the GCC. We approach the performance characterization problem in the following way: the performance of the optimal estimator in the presence of large errors is related to the probability that the likelihood function gives an estimate which is far removed from the true time delay. This probability, in turn, admits a representation in terms of level up-crossing probabilities. In Chapter 4, we present a general theorem on the asymptotic behavior of the level up-crossings of a non-stationary random process. Then, based on the results of Chapter 4, a Poisson model is invoked as an approximation to the level crossings and a numerical study of the large error probability and the global variance is undertaken for lowpass and bandpass signal spectra.

Errata

p7:   Line 4 should read " the variance of $\hat{R}_{12}$ "

p11:   Eqs. (20) and (21): " $\int_{0}^{T}$ " replaces " $\int_{-T}^{T}$ ".

p16:   Line 7 and 8: delete " (see Eqns. ...) ".

p71:   Eq. 1: " ∃ " replaces " E ".

p85:   Eq. (43): " - " in exponential argument should be " + ".

p163:   Fig. 2.1: Shift ordinates, " $2\pi/B$, $-2\pi/B$, ... ,$6\pi/B$, $-6\pi/B$ " to adjacent zero crossings of auto-correlation function.

p170:   Figure caption should read " ... Bandpass ... " not " ... Banpass ... ".

Table of Contents

List of Figures

Chapter 1
Introduction

Time delays are inherent to many important problems in a variety of scientific and engineering disciplines. In observational astronomy, very long baseline interferometry (VLBI) uses the relative delays between multiple measurements, taken at widely separated sensors, to compute the positions of heavenly bodies. In radio communications, increases in the received signal energy are possible through coherent (time-aligned) combination of the outputs of a multi-sensor antenna array. As an example from medical research, studies in evoked potentials involve computing the reaction time of a subject to different forms of stimuli. Essentially, in any situation where measurements are taken at multiple points in space in order to resolve or utilize certain characteristics of a propagating signal, the inter-sensor propagation time is of prime importance.

For the two sensor, linear inter-sensor channel case, cross- correlation of the sensor outputs produces a global peak at the true delay. However, the presence of significant random sensor noise makes it essential to pre-process the outputs, prior to correlation, in order to obtain acceptable peak type estimates of the delay parameter. The susceptability of correlation peak estimates to various factors such as: unknown signal statistics, low signal bandwidth, high center frequency and small observation time, is well documented in the literature. In this thesis we concentrate on: 1). developing cross-correlation techniques which are less sensitive to the above factors, and 2). better characterizing the (large error) performance of peak detection type correlation estimates relative to the unfavorable conditions above. The main results in these directions are organized into four chapters (Chapters 2,3 and 6 deal with 1)., Chapter 5 deals with 2).). The remaining chapter, Chapter 4, contains results which are supportive to Chapter 5. Chapters 2,3,4,5 and 6 are briefly summarized below.

Chapter 2 sets up the formal problem definition. Then a unifying interpretation of several of the classical Generalized Cross Correlation (GCC) methods is given in terms of signal-to-noise ratio (SNR). The main result is the development of a new GCC motivated from Least Mean Square Error (LMSE) filtering of the sensors. A first step towards obtaining robust estimates of time delay, under uncertainty on the signal and noise spectra, is acheived by implementing robust Wiener filters in place of the global LMSE filters for the signal. Finally a numerical evaluation of performance is undertaken to compare the new techniques to some of the well known GCCs.

In Chapter 3, an alternate approach is proposed: one obtains the time delay estimate by means of an approximation to the center of symmetry of the GCC rather by the conventional peak detection method. A class of center of symmetry estimators (CSE) is defined using the Fourier coefficients of the GCC, and an approximate expression for the estimator variance is derived. An optimal (minimum variance) CSE is then obtained which is compared to the optimal (minimum variance) GCC by simulation. The results of the simulation indicate that for the difficult case of narrowband signals, the optimal CSE has much better performance than the GCC for low SNR, and nearly identical performance for high SNR.

While Chapter 4 is not directly related to time delay estimation, it provides the theoretical background for Chapter 5, which is concerned with more widely applicable GCC variance approximations than those currently available in the literature. Chapter 4 contains material on the asymptotic distribution of level crossing probabilities. Several authors have studied the weak convergence of the number of level crossings, $N(t)$, of a given stationary random process over an interval $[0,t]$, to a Poisson random count process, as the level is allowed to become large and the length of the interval is suitably normalized. In Chapter 4

results are obtained for the distribution of level up-crossings by a non-stationary random process. It is established that, under certain regularity conditions, the asymptotic distribution is again Poisson, but with an inhomogeneous (time varying) rate parameter. This is accomplished by showing that the up-crossing probability obeys an asymptotically linear, first order integral equation.

In Chapter 5 a level crossing approach is developed to model the large error performance of the GCC. In the past, large error variance expressions were developed for the large bandwidth-time product (BT) case, which involve computing the probabilities of level exceedance by a finite, but large, number of time samples along the GCC trajectory. Essential to these previous methods is the assumption that the sidelobe activity of the GCC is negligible, a tenuous assumption for low BT or highly narrowband signal spectra. In Chapter 5 the expected number of up-crossings of a level by the GCC yields insights into the large error performance of the GCC for situations where reliable indicators could only be provided by experiment. Motivated by the results of Chapter 4, a Poisson approximation to the large error variance is derived. Numerical studies of the Poisson approximation reveal several effects on performance of highly narrowband signals. Some of these effects match the behavior of the Ziv-Zakai Lower Bound (ZZLB) on the variance of the optimal estimate, but were not detected by other large error approximations reported in the literature.

Finally Chapter 6 informally explores two new and interesting techniques of performing time delay estimation. One of these is developed specifically for highly oscillatory periodic signals. It is motivated by the behavior of lower bounds (the ZZLB and others) on the variance of the optimal estimator. The resultant structure for the estimator is a hybrid GCC processor which acts on the observations and their envelopes to obtain an optimal estimate. A constrained optimization of the local (small error) and the global (large error) variance specifies a GCC

which, on the one hand, implements the locally optimal estimator, and on the other hand, passes the locally optimal estimate through a tapped delay line (equalization) to suppress its large error-prone sidelobes. The second technique investigated in this chapter uses a functional transformation on the cross-correlation which allows us to replace the inherently non-linear time delay estimation problem with a linear estimation problem (observations = delay + noise) under certain circumstances. Finally some optimal and sub-optimal processing structures are presented, based on the above linearization. A preliminary analysis of performance concludes the chapter.

Chapter 2
The Wiener Processor for Time Delay

I. Introduction

In this chapter a general form of the time delay estimation problem of interest will be stated. In Section II, cross-correlation type processors will be motivated and the Generalized Cross-Correlator (GCC) will be defined. A correlation domain signal-to- noise ratio (SNR) approach will be adopted to derive three of the classical "optimal" GCCs in a unified manner. Although the concept of using SNR as a GCC optimization criterion is not new (two of these optimal GCCs were derived in this way by their respective authors), the *Hannan–Thomson* GCC (*HT*) has not previously been considered in this light. In chapter 6 one of the SNRs defined here, the one associated with the optimization yielding the *HT*, will be shown to have a special significance for the theoretical performance of the globally optimal (minimum variance) estimator. We conclude the introductory part of this chapter with a discussion of an "ad hoc" or intuitive GCC, the *Smoothed Coherence Transform* (*SCOT*).

In Section III, the main result of the chapter is presented, the derivation of the *Wiener Processor* (*WP*). This derivation is obtained by focusing on the observation domain, as opposed to the cross-correlation domain. A numerical study is presented in Section V which indicates good performance of the *WP* relative to the four other GCCs considered here.

In general the optimum processors are very sensitive to deviations from the assumed signal and noise characteristics. By way of contrast, the *SCOT* appears to be more robust to these deviations from the nominal model, as has been indicated by experimental results. However, these latter processors can have very poor performance at the nominal point.

In Section IV of this chapter, an implementation of the WP as a "robust" processor for spectral uncertainty (RWP) is discussed. The numerical results of Section V indicate a RWP performance gain, over the classical GCCs mentioned above, at the least favorable point in an ϵ-contaminated uncertainty class of spectra. Hence, for the cases under consideration, the WP and RWP compare favorably to existing GCC methods of time delay estimation.

II. Problem Statement and Background

We first consider a system model generating the observations in Fig. 1. We observe Gaussian, ergodic, wide-sense stationary processes

$$x_1(t) = s(t) + n_1(t) \tag{1}$$

$$x_2(t) = s_o(t) + n_2(t)$$

over a time interval $t \in [0, T]$. The noises $n_1(t)$ and $n_2(t)$ are uncorrelated broadband Gaussian proceses, $s(t)$ is a Gaussian signal and $s_o(t) = c(t) * s(t)$ is a filtered version of $s(t)$. In this chapter $c(t)$ is a linear, time invariant channel having a transfer function $C(\omega)$, with unknown linear phase, so that $s_o(t)$ is a delayed, but possibly distorted, version of $s(t)$. This delay, D, is *a priori* known to lie within the time interval $[-D_m, D_m]$. Furthermore, we assume that the noises are uncorrelated with the signal and that the observation time, T, is much greater than both the correlation time, T_c, of $x_1(t)$ and $x_2(t)$, and D_m. The object is then to estimate the time delay, D, associated with the channel.

We define the sample (averaged) cross-correlation

$$\hat{R}_{12}(\tau) = \frac{1}{T} \sum_{k=1}^{m} \int_0^T x_{1k}(\sigma) x_{2k}(\sigma+\tau) d\sigma \tag{2}$$

where $x_{ik}(t)$ is a time truncated version of $x_i(t)$ for $i = 1, 2$

$$x_{ik}(t) = \begin{cases} x_i(t), & t \in [(k-1)T/m, kT/m] \\ 0, & otherwise \end{cases} \qquad (3)$$

where it is assumed that $T/m \gg T_c$ and $T/m \gg D_m$. When $m = 1$, in (2), the sample cross-correlation is called a coherent estimate of the cross-correlation function R_{12}. If T is sufficiently large one can choose $m > 1$ in order to reduce the variance of $R_{\hat{1}2}$. This latter procedure gives the so-called incoherent cross-correlation estimate. For simplicity of presentation, $m = 1$ in the remainder of this chapter. Alternately (2) can be expressed in the frequency domain (for $T \gg T_c$)

$$\hat{R}_{12}(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{G}_{12}(\omega) e^{j\omega\tau} d\omega \ . \qquad (4)$$

Here $\hat{G}_{12} = \frac{1}{T} X_1^* X_2$, where X_1 and X_2 are the finite time Fourier transforms of x_1 and x_2. For large observation time, $\hat{R}_{12}(\tau)$ is a good approximation to the true cross-correlation function which has a global peak at D. In fact if $c(t)$ is pure delay and $s(t)$ is white, the cross-correlation function is a delta function at the true delay. For finite observation time we can decompose $\hat{R}_{12}(\tau)$ into the sum of four terms

$$\hat{R}_{12}(\tau) = c(\tau) * \hat{R}_{ss}(\tau) + c(\tau) * \hat{R}_{n_1 s}(\tau) + \hat{R}_{sn_2}(\tau) + \hat{R}_{n_1 n_2}(\tau) \qquad (5)$$

where "*" denotes convolution. Here $\hat{R}_{ss}(\tau)$ is an estimate of the signal auto-correlation function $R_{ss}(\tau)$, and $\hat{R}_{n_1 s}(\tau)$, $\hat{R}_{sn_2}(\tau)$ and $\hat{R}_{n_1 n_2}(\tau)$ are estimates of the cross-correlation between the respective signal and noise terms in the observations. In the limit ergodicity guarantees that the sample cross-correlation converges to $c(\tau) * R_{ss}(\tau)$, which displays an absolute maximum at D. Thus, it is the last three terms in Eq. (5) which constitute zero mean disturbances affecting the peak resolution of the first term. This suggests prefiltering the sample cross-

correlation with a filter $W(\omega)$ to obtain better resolution of the peak at D. To avoid the introduction of an unnecessary bias into the peak location we confine our attention to zero phase $W(\omega)$. The resultant filtered sample cross-correlation is then searched over the *a priori* region $[-D_m, D_m]$ for the global peak, which yields the time delay estimate, $\hat{D}$. This scheme is referred to as the generalized cross-correlation method or the Generalized Cross-Correlator (G.C.C.) and is illustrated in Fig. 2. We denote the G.C.C. output waveform $R_{12}^g(\tau)$. Therefore we have

$$R^g(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{G}_{12}(\omega) W(\omega) e^{j\omega\tau} d\omega \ . \tag{6}$$

$$\hat{D} = \underset{\tau \in [-D_m, D_m]}{argmax} R^g(\tau)$$

When $W(\omega)$ is unity the resulting G.C.C. is called the simple cross-correlator (C.C.). Considering the first term in Eq. (5) as a "signal" in additive noise, classical optimal filtering theory can be applied to derive filters $W(\omega)$ which maximize signal-to-noise ratio.

Letting the last three terms of Eqn (5) be characterized as "noise" we can define a signal-to-noise ratio at the output of the GCC as the magnitude squared of the global peak of the "signal" term divided by the power of the "noise" which generates false peaks in $R_{12}^g(\tau)$. We will denote this SNR_1. For a sufficiently broadband signal, $s(t)$, the variance of the cross-correlation estimate, Eq. (2), outside of the immediate vicinity of the true delay is given by

$$var(\hat{R}_{12}(\tau)) = \frac{1}{T2\pi} \int_{-\infty}^{\infty} G_{11}(\omega) G_{22}(\omega) d\omega \tag{7}$$

and the variance of $\hat{R}_{ss}$ by

$$var(\hat{R}_{ss}(\tau)) = \frac{1}{T2\pi} \int_{-\infty}^{\infty} |G_{12}(\omega)|^2 d\omega \tag{8}$$

for $\tau \neq D$ [12]. $G_{11}(\omega)$ and $G_{22}(\omega)$ are the power spectral densities of the observations $x_1(t)$ and $x_2(t)$ respectively and $G_{12}(\omega)$ is the cross-spectrum. Using the above results it is straightforward to derive the cross-correlation noise power which is given by

$$\sigma_n^2(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} G_{11}(\omega) G_{22}(\omega)(1-|\gamma_{12}(\omega)|^2)|W(\omega)|^2 \, d\omega \quad , \qquad \tau \neq D \qquad (9)$$

$|\gamma_{12}(\omega)|^2$ is the magnitude coherency squared

$$|\gamma_{12}(\omega)|^2 = \frac{|G_{12}(\omega)|^2}{G_{11}(\omega) G_{22}(\omega)} \quad . \qquad (10)$$

Then from the defining relation

$$SNR_1 = \frac{[E\{R^g(\tau)|_{\tau=D}\}]^2}{\sigma_n^2} \qquad (11)$$

we obtain

$$SNR_1 = \frac{[\frac{1}{2\pi} \int_{-\infty}^{\infty} |G_{12}(\omega)| W(\omega) d\omega]^2}{\frac{1}{2\pi} \int_{-\infty}^{\infty} G_{11}(\omega) G_{22}(\omega)(1-|\gamma_{12}(\omega)|^2)|W(\omega)|^2 d\omega} \qquad (12)$$

The maximum is obtained through the Schwarz inequality and yields the H.T. processor for the pure delay channel. The same result is derived in [4] as the result of minimizing the local variance of the delay estimate over the entire G.C.C. class, and in [1] as the result of maximum likelihood estimation. The filter is

$$W_{H.T.}(\omega) = \frac{1}{|G_{12}(\omega)|} \frac{|\gamma_{12}(\omega)|^2}{1-|\gamma_{12}(\omega)|^2} \quad . \qquad (13)$$

Neglecting the effect of the signal and noise cross terms , $c(\tau) * \hat{R}_{sn_1}(\tau)$ and $\hat{R}_{sn_2}(\tau)$ in Eq. (5), gives another characterization of the noise in the cross-

correlation domain. With this definition of noise another signal-to-noise ratio is defined in [6], SNR_2, which is shown, for pure delay, to be maximized by the *Eckart* processor $W_{EK}(\omega)$

$$SNR_2 = \frac{[\frac{1}{2\pi} \int\limits_{-\infty}^{\infty} |G_{12}(\omega)| W(\omega) d\omega]^2}{\frac{1}{2\pi} \int\limits_{-\infty}^{\infty} G_{n_1}(\omega) G_{n_2}(\omega) |W(\omega)|^2 d\omega} \qquad (14)$$

$$W_{EK}(\omega) = \frac{|G_{12}(\omega)|}{G_{n_1}(\omega) G_{n_2}(\omega)} \qquad (15)$$

where $G_{n_1}(\omega)$ and $G_{n_2}(\omega)$ are the auto-spectra of the noises $n_1(t)$ and $n_2(t)$ respectively. Note that in terms of the spectra of the observables $x_1(t)$ and $x_2(t)$, the filter takes the form

$$W_{EK}(\omega) = \frac{|G_{12}(\omega)|}{(G_{11}(\omega) - |G_{12}(\omega)|)(G_{22}(\omega) - |G_{12}(\omega)|)}$$

Hassab and Boucher [2] take the approach of maximizing a signal-to-noise ratio, SNR_3, defined as the ratio of the expected peak energy at the true delay to the total statistical variation of the output of the G.C.C. This, in a sense, lumps the "signal", $c(\tau) * \hat{R}_{ss}(\tau)$, variation into the noise terms and yields the *H.B.* filter $W_{H.B.}(\omega)$

$$SNR_3 = \frac{[\frac{1}{2\pi} \int\limits_{-\infty}^{\infty} |G_{12}(\omega)| W(\omega) d\omega]^2}{\frac{1}{2\pi} \int\limits_{-\infty}^{\infty} G_{11}(\omega) G_{22}(\omega) |W(\omega)|^2 d\omega} \qquad (16)$$

$$W_{H.B.}(\omega) = \frac{|G_{12}(\omega)|}{G_{11}(\omega) G_{22}(\omega)} \qquad (17)$$

The *H.B.* is similar to the *SCOT* introduced by Carter *et al* [5] in that, for

highly dynamic spectra, in addition to suppressing the cross-spectral estimate in ω-regions of low signal-to-noise ratio, high signal-to-noise ratio regions are also suppressed in an attempt to reject strong tones in the observations.

Note that the above performance criteria impose equal penalties on small and large errors. That is the location of the false peak in the GCC output exerts no influence on the signal-to-noise ratios defined in Eqns. (11), (14) and (16). Therefore one can only rely on these criteria if the signal-to-noise ratio is sufficiently high to guarantee a low probability of large error. The behavior of this probability as a function of signal-to-noise ratio, observation time and signal bandwidth is investigated elsewhere [7,13].

III. The Wiener Processor

Here a different approach is taken to derive an optimal filter. We deal directly with the quantities in the observation time domain (i.e. Fig. 1). The procedure is motivated by the following argument. If we knew the signal $s(t)$ and the filtered version $s_o(t)$ exactly, then, from the linearity of the phase of the channel, the time delay could be estimated exactly by detecting the peak of the sample cross-correlation of $s(t)$ and $s_o(t)$. Therefore we simply try to estimate the signal $s(t)$ as best we can from the observations $x_1(t)$ and the channel output signal $s_o(t)$ from $x_2(t)$ by minimizing the mean square errors

$$E\{(s(t) - \hat{s}(t))^2\} = \min \tag{18}$$
$$E\{(s_o(t) - \hat{s}_o(t))^2\} = \min \tag{19}$$

where

$$\hat{s}(t) = \int\limits_{-T}^{T} x_1(\sigma)h_1(t-\sigma)d\sigma \tag{20}$$

$$\hat{s}_o(t) = \int\limits_{-T}^{T} x_2(\sigma)h_2(t-\sigma)d\sigma \ . \tag{21}$$

The above procedure is illustrated in Fig. 3. Given the channel characteristic $C(\omega)$ the solutions to Eq. (18) and (19) are the Wiener filters $H_1(\omega)$ and $H_2(\omega)$

$$H_1(\omega) = \frac{G_{ss}(\omega)}{G_{ss}(\omega) + G_{n_1 n_2}(\omega)} \tag{22}$$

$$H_2(\omega) = \frac{G_{ss}(\omega)|C(\omega)|^2}{G_{ss}(\omega)|C(\omega)|^2 + G_{n_2 n_2}(\omega)} . \tag{23}$$

Noting that $G_{12}(\omega) = C(\omega)G_{ss}(\omega)$ we can express the above filters in terms of the quantities derived from the observables

$$H_1(\omega) = \frac{1}{C(\omega)} \frac{G_{12}(\omega)}{G_{11}(\omega)} \tag{24}$$

$$H_2(\omega) = C^*(\omega) \frac{G_{12}(\omega)}{G_{22}(\omega)} . \tag{25}$$

where $C^*(\omega)$ is the complex conjugate of $C(\omega)$.

With these filters the sample cross-correlation of the least mean square error estimates of $s(t)$ and $s_o(t)$ yields the estimate of the cross correlation function

$$R^{WP}(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{1}{T} \hat{S}^*(\omega)\hat{S}_o(\omega)e^{j\omega\tau}d\omega \tag{26}$$

where

$$\hat{S}(\omega) = H_1(\omega)X_1(\omega) \tag{27}$$

$$\hat{S}_o(\omega) = H_2(\omega)X_2(\omega) \tag{28}$$

Regrouping terms in (26) we obtain

$$R^{WP}(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{G}_{12}(\omega) \frac{|G_{12}(\omega)|^2}{G_{11}(\omega)G_{22}(\omega)} e^{j\omega\tau}d\omega \tag{29}$$

Comparing Eq. (29) with Eq. (6) we have the result that the *W.P.* is equivalent to using a Generalized Cross-Correlator with the filter $W(\omega)$ equal to the magnitude coherency squared.

It should be emphasized that even though the Wiener filters H_1 and H_2 involve the knowledge of the channel $C(\omega)$ itself the G.C.C. equivalent processor does not impose this requirement. In fact, as far as the cross-correlation estimate of time delay is concerned, the actual channel is immaterial to the peak detection procedure in the cross-correlation domain. Hence the Wiener filter implementation (Fig. 3) with $C(\omega)$ arbitrarily set to unity in Eqns. (24) and (25) is equivalent to any other choice of $C(\omega)$ for the time delay estimation problem.

From Eq. (7) the variance of the cross-spectrum estimate $\hat{G}_{12}(\omega)$ is proportional to the product of the auto-spectra of the observations $G_{11}(\omega)G_{22}(\omega)$. Fix the sample auto-correlation $\hat{R}_{ss}(\tau)$ in Eq. (5). Then the definition of "additive noise" leading to the signal-to-noise ratio SNR_1, Eq. (9), yields the interpretation of $1 - |\gamma_{12}(\omega)|^2$ as a measure of the cross-spectral estimator variance about the "desired signal" $c(\tau) * \hat{R}_{ss}(\tau)$. Thus the W.P. de-emphasizes those ω-regions where the sample cross-spectrum is likely to be a highly inaccurate estimate of the true cross-spectrum. This is not surprising given the *raison d' être* of the W.P. which is to estimate accurately the smoothed sample auto-correlation, $c(\tau) * \hat{R}_{ss}(\tau)$.

The W.P. does not of course maximize the signal-to-noise ratio in general. If we examine the optimal processor for SNR_1, the H.T. (Eq. (13)), we see that it has the additional ability to overemphasize as well as to de-emphasize the cross-spectral estimate according to the function $|\gamma_{12}(\omega)|^2/(1 - |\gamma_{12}(\omega)|^2)$. (Actually in [4] the above function is shown to be inversely proportional to the variance of the phase estimate $\hat{G}_{12}(\omega)/|\hat{G}_{12}(\omega)|$ with respect to the true phase of the cross-spectrum). However, in situations where the coherence is low, and the signal spectrum nearly flat, the H.T. and the W.P. are virtually identical and exhibit identical performance (Eq. (13) becomes proportional to $|\gamma_{12}|^2$).

It is also observed that the *W.P.* is equivalent to the *H.B.* for nearly flat signal spectra and also to the *Eckart* if we add a low signal-to-noise ratio condition

$$W_{H.B.}(\omega) = \frac{1}{|G_{12}(\omega)|} \; |\gamma_{12}(\omega)|^2 = \frac{1}{|G_{12}(\omega)|} \; W_{W.P.}(\omega) \tag{30}$$

$$W_{EK}(\omega) = \frac{G_{ss}(\omega)}{G_{n_1}(\omega)G_{n_2}(\omega)} \; \approx \; \frac{1}{|G_{12}(\omega)|} \; |\gamma_{12}(\omega)|^2 \tag{31}$$

The above signal-to-noise ratio condition is that $G_{ss}(\omega)$ be uniformly small as compared to $G_{n_1}(\omega)$ and $G_{n_2}(\omega)$.

IV. The Robust Wiener Processor For Unknown Spectra

The optimal GCC's all require knowledge of the signal and noise spectra underlying the observations. When the spectral quantities used in the filter function for the GCC do not match the true spectra there is a consequent deterioration in performance. Two approaches to the problem of unknown spectra are of interest. We either estimate the spectra and substitute the estimates into the aforementioned filters (totally unknown spectra) or we search for a robust solution over a range of spectra perturbed from some nominal point (partially unknown spectra).

For the estimation approach the sensitivity of the GCC filter to small deviations in the estimated spectra may be an important consideration. As applied to the *H.T.*, the substitution method yields a procedure which weights the phase of the sample cross-spectrum, $\hat{G}_{12}(\omega)$, with the function $\hat{W}_{H.T.}(\omega) = |\hat{\gamma}_{12}(\omega)|^2/(1 - |\hat{\gamma}_{12}(\omega)|^2)$, $|\hat{\gamma}_{12}(\omega)|^2$ a magnitude coherency squared estimate. A simple local analysis of the estimation error associated with $\hat{W}_{H.T.}$ yields the variance

$$var(\hat{W}_{H.T.}(\omega)) \approx \frac{1}{(1- |\gamma_{12}(\omega)|^2)^4} \; var(|\hat{\gamma}_{12}(\omega)|^2) \; . \tag{32}$$

$\hat{W}_{H.T.}$ may critically underweight the phase estimate over frequencies where $|\gamma_{12}|^2$ is high, that is where the phase estimate is apt to be the most accurate. On the other hand substitution of $|\overset{\wedge}{\gamma}_{12}|^2$ for the W.P. gives only as much error as the estimation error of $|\overset{\wedge}{\gamma}_{12}|^2$ itself. An improved filter estimate could translate into improved performance of the time delay estimate. Naturally these comments must be verified through a future simulation study.

In practice the spectra may be only partially unknown and a different strategy can be used to design the GCC. This is the "robust" approach which has been applied to classical matched and Wiener filtering with some success [8-10]. The resultant filters are robust in the maximin sense, e.g. the filter maximizes the minimum output signal-to-noise ratio as the spectra are allowed to vary over their regions of uncertainty. Usually, one first finds the signal and noise spectral pair which gives the least favorable performance, if it exists. Then one optimizes the filter for the least favorable pair, hence the name maximin filter. For a precise formulation of the robust approach see [14] and [8-10].

To our knowledge, no results are known concerning the solution of the robust time delay estimation problem. Short of this the only known published result in maximin filters for time delay is that of Kassam and Hussaini [11] for the pure delay case. In [11] they used the fact that the Eckart processor maximizes a classically defined signal-to-noise ratio (See Eq. (14)) to relate the filtering problem to robust hypothesis testing. This is achievable only by associating uncertainty classes with the spectral product $G_{n_1}(\omega)G_{n_2}(\omega)$, rather than with the individual noise spectra themselves.

For the pure delay channel an alternate approach to combatting against poor performance with uncertain spectra is suggested by the recent work in robust Wiener filtering, [8],[9], when applied to the W.P. With regard to the original formulation of the W.P., we can replace the least mean square estimates of

the channel input and the channel output by the robust least mean square estimates of $s(t)$ and $s(t-D)$ under uncertainty in the signal and noises. Specifically we assume that the signal spectrum $G_{ss}(\omega)$ belongs to the spectral class $\{\sigma\}$, and that the noise spectra $G_{n_1}(\omega)$ and $G_{n_2}(\omega)$ belong to the spectral classes $\{\eta_1\}$ and $\{\eta_2\}$ respectively. Then we solve for the least favorable pairs for Wiener filtering $\{G_{ss}^{lf}(\omega), G_{n_1}^{lf}(\omega)\}$ and $\{G_{ss}^{lf}(\omega), G_{n_2}^{lf}(\omega)\}$ over the product classes $\{\sigma \times \eta_1\}$ and $\{\sigma \times \eta_2\}$ which yield the robust Wiener filters H_1^R and H_2^R (see Eqns. (2.3) and (2.4)). Finally we implement these filters in the cross-correlation domain as a G.C.C., a scheme which we will call the *Robust Wiener Processor* or the *R.W.P.*

V. Numerical Comparisons

At the present time no simulation results concerning the experimental performance of the *W.P.* and *R.W.P.* as opposed to the other G.C.C.'s are available. In their absence a preliminary investigation of the relative merits of the above processors was performed based on the various signal-to-noise ratio criteria defined in Section II, for some specific observation spectra and for the pure delay channel.

Example 1

Figs. 6 through 9 show the relative performance of the *H.T.,H.B.,Eckart*, *SCOT*, and *C.C.* under the criteria SNR_1, SNR_2, SNR_3 and local variance, var_L, of the time delay estimate [3], for a third order Markov signal in first order Markov noises with the noise 3dB bandwidth a factor of ten greater than that of the signal (see Figs. 4 and 5). These spectra were chosen for their tail behavior to avoid certain degeneracies in the local variance criterion. The interesting thing to note is that under SNR_1 and SNR_2 the *W.P.* exhibits better performance

than all of the other sub-optimum G.C.C.'s for that particular definition of SNR. Under SNR_3 it is a close second next to the *M.L.E.* In fact, under the criterion SNR_1, performance of the *W.P.* is virtually identical to the optimal *H.T.* processor. Although the local variance ranks the *W.P.* behind the *H.T.*, *H.B.* and *Eckart* (see Fig. 9), it only marginally disfavors the *W.P.* at low signal-to-noise ratios (In Fig. 9 the *SCOT* and *C.C.* processors have orders of magnitude worse performance than the *W.P.* and are off scale).

Example 2

Here the performance of the *R.W.P.*, *W.P.* and other G.C.C.'s are compared using SNR_1 for the ϵ-contaminated uncertainty class on the specific spectra in the example outlined in Kassam and Lim's paper on Robust Wiener filtering [9]. Specifically, under the nominal assumption , at each sensor we have a signal with the flat bandlimited spectrum $G_{ss}^o(\omega)$ in first order Markov noise with the spectrum $G_n^o(\omega)$, where the signal and noises are of comparable bandwidths (see Figs. 10 and 11). The uncertainty on the signal and noise spectra are modeled as the ϵ-mixtures

$$(1-\epsilon_1)G_{ss}^o(\omega) + \epsilon_1 G_{ss}'(\omega) \tag{33}$$

and

$$(1-\epsilon_2)G_n^o(\omega) + \epsilon_2 G_n'(\omega) \tag{34}$$

respectively with $G_{ss}'(\omega)$ and $G_n'(\omega)$ arbitrary spectra having the same mass as the nominal and ϵ_1 and ϵ_2 lying in the interval [0,1]. When $\epsilon_1 = 0.2$ and $\epsilon_2 = 0.1$ the least favorable signal and noise spectra are plotted in Figs. 12 and 13 respectively. This corresponds to the case where one may have more confidence in the nominal noise than in the nominal signal. The least favorable spectra for this example illustrate a typical attribute of least favorables in that the worst performance of a Wiener filter occurs when the signal masquerades as the

noise and vice-versa, i.e. when we get a minimum separation of hypotheses concerning the presence or absence of the signal within the uncertainty classes (33) and (34). Fig. 14 shows the relative performance for the nominal spectra and Fig. 15 the performance for the least favorable signal and noise spectra for Wiener filtering. Looking at the nominal case we note that the use of the *R.W.P.* entails a loss of about 3dB at low SNR (below about 0db) over the optimal for the least favorable pair. However when the true signal and noise spectra are least favorable for Wiener filtering the *R.W.P.* displays uniformly better relative performance, gaining about 3db over the other processors at low signal-to-noise ratios. Note that the pair in Figs. 12 and 13 is not necessarily the least favorable pair for *H.T.* filtering so that no conclusive result is indicated here. However, Fig. 15 does suggest that, at least for some spectra in the above uncertainty class, we can expect better performance with the *R.W.P.* than with the optimal scheme for the nominal spectra. Therefore the *WP* and *RWP* instrumentations may be viable alternatives to existing time delay estimation schemes. However, experimental investigations of these advantages will prove to be the ultimate yardstick of performance.

REFERENCES

Chapter 2

1. C.H. Knapp and G.C. Carter, "The Generalized Correlation Method for Estimation of Time Delay," *IEEE Trans. Acoust., Speech, Signal Processing,* Vol. ASSP-24, pp. 320-327, August 1976.

2. J.C. Hassab and R.E. Boucher, "Optimum Estimation of Time Delay by a Generalized Correlator," *IEEE Trans. Acoust., Speech, Signal Processing,* Vol. ASSP-27, pp. 373-380, August 1979.

3. V.H. MacDonald and P.M. Schultheiss, "Optimum Passive Bearing Estimation in a Spatially Incoherent Noise Environment," *J. Acoust. Soc. Amer.,* Vol. 46, No. 1, pp. 37-43, 1969.

4. E.J. Hannan and P.J. Thomson, "Estimating Group Delay," *Biometrika,* Vol. 60, pp. 241-253, 1973.

5. G.C. Carter, A.H. Nuttall, and P.G. Cable, "The Smoothed Coherence Transform," *Proc. IEEE,* Vol. 61, No. 10, pp. 1497-1498, 1973.

6. C. Eckart, "Optimal Rectifier Systems for the Detection of Steady Signals," University of California, Scripps Inst. Oceanography, Marine Physics Lab., Report S10 12692, S10 Ref. 52-11, 1952.

7. A.J. Weiss and E. Weinstein, "Fundamental Limitations in Passive Time Delay Estimation-Part I: Narrow-Band Systems," *IEEE Trans. Acoust., Speech, Signal Processing,* Vol. ASSP-31, No. 2, pp. 472-485, April 1983.

8. S.A. Kassam and T.L. Lim, "Robust Wiener Filters," *J. Franklin Institute,* Vol. 304, pp. 171-185, 1977.

9. H.V. Poor, "On Robust Wiener Filtering," *IEEE Trans. Automat. Contr.,* Vol. 25, No. 3, pp. 531-536, June 1980.

10. H.V. Poor, "Robust Matched Filters," *IEEE Trans. on Inform. Theory,* Vol. IT-29, No. 5, pp 677-687, September 1983.

11. E.K. Al-Hussaini and S. Kassam, "Robust Filters for Time Delay Estimation Problems," *Proc. of the Sixteenth Annual Conf. on ISS,* Princeton University, pp. 540-545, March 1982.

12. J.S. Bendat and A.G. Piersol *Random Data: Analysis and Measurement Procedures,* Wiley, Inter-science Series, 1971.

13. J. P. Ianniello, "Time Delay Estimation Via Cross-Correlation in the Presence of Large Estimation Errors," *IEEE Trans. on ASSP,* Vol. ASSP-30, No. 6, pp. 998-1003, Dec. 1982.

14. E.L. Lehmann, *Testing Statistical Hypotheses,* Wiley, N.Y., 1959.
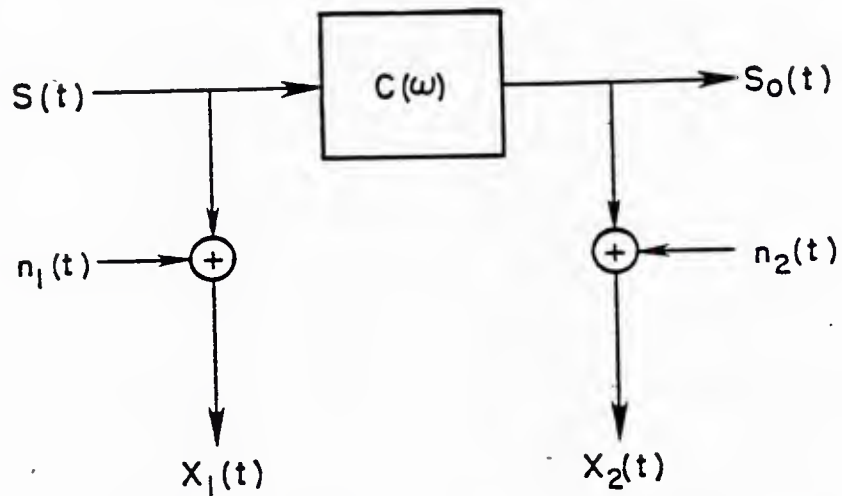
Fig. 1: Model governing the observations $x_1(t)$ and $x_2(t)$
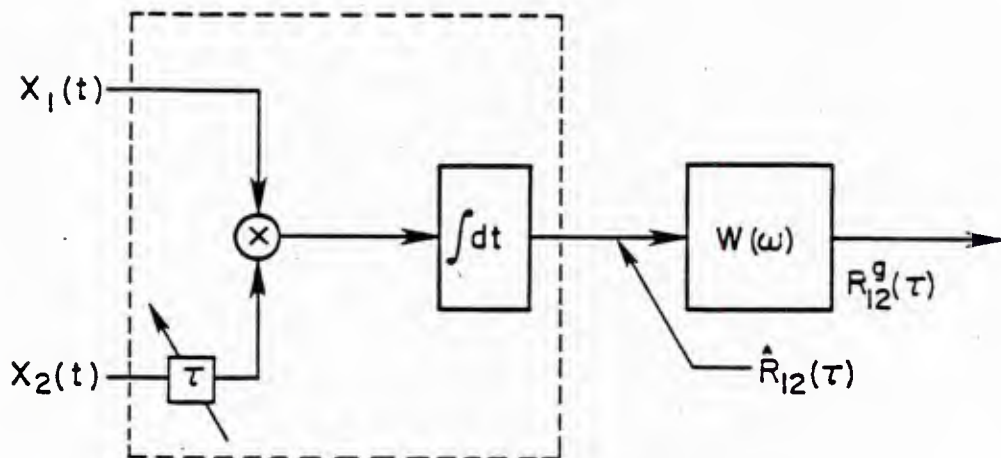
Fig. 2: GCC implemented as one postfilter $W(w)$
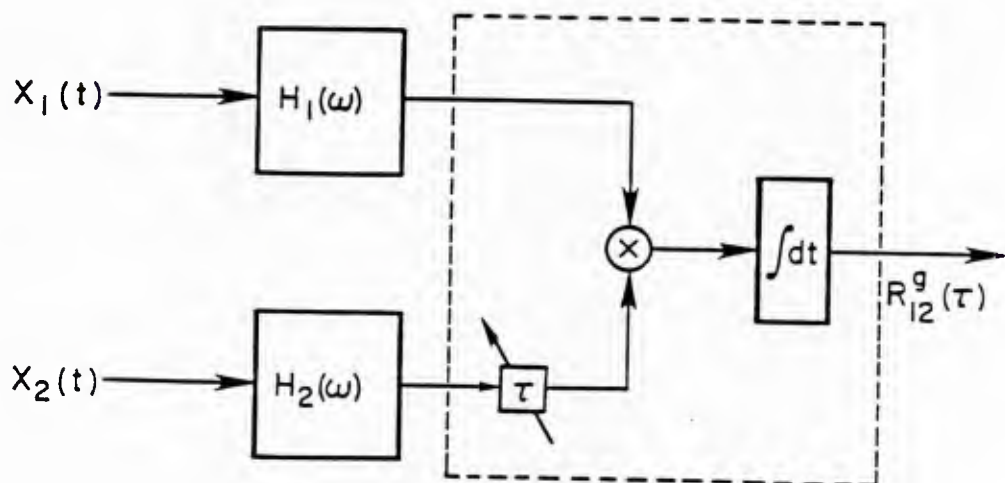
Fig. 3: GCC implemented as two prefilters $H_1(w)$ and $H_2(w)$

Third order Markov signal spectrum for Ex. 1
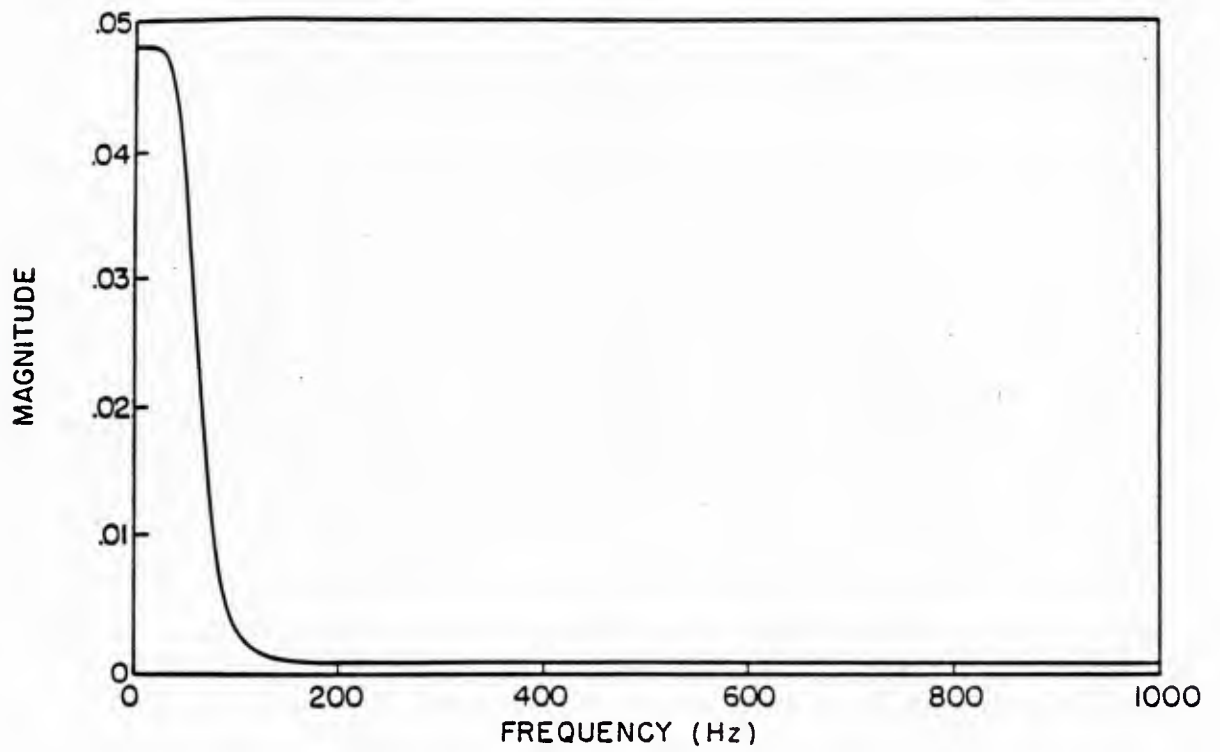
Fig. 4

First order Markov noise spectrum for Ex. 1
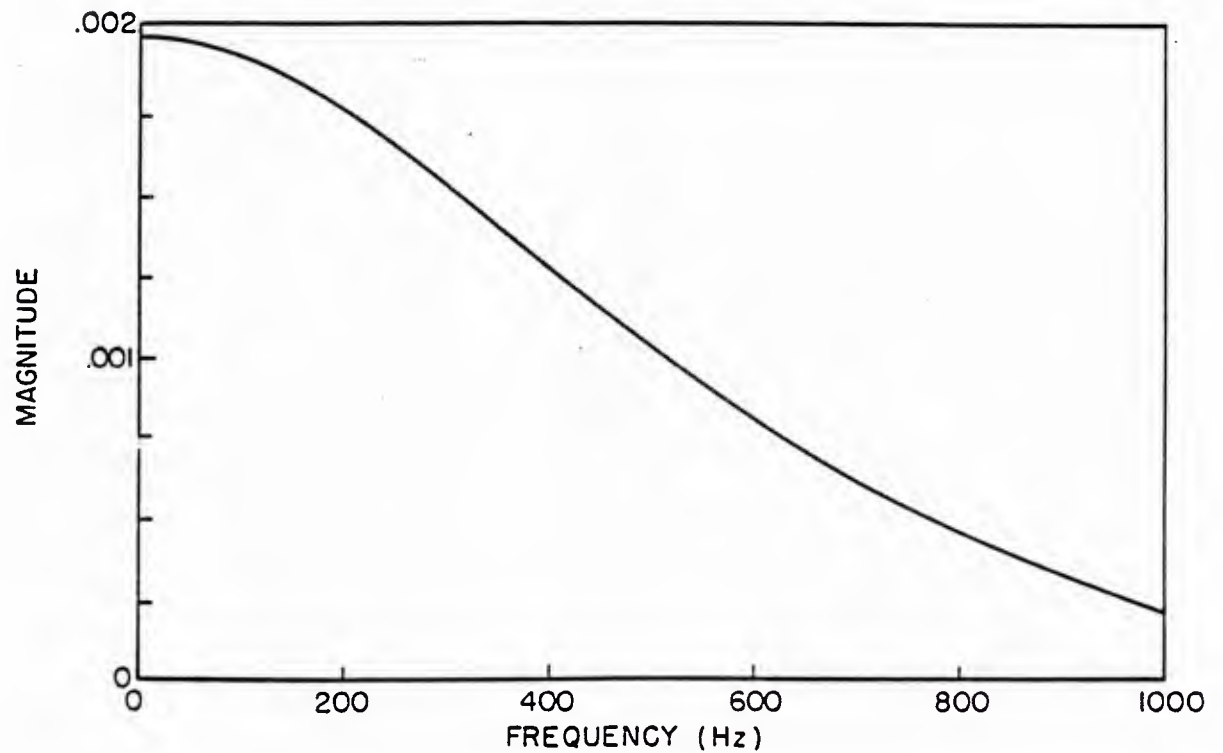
Fig. 5

Performance of various processors under SNR_1
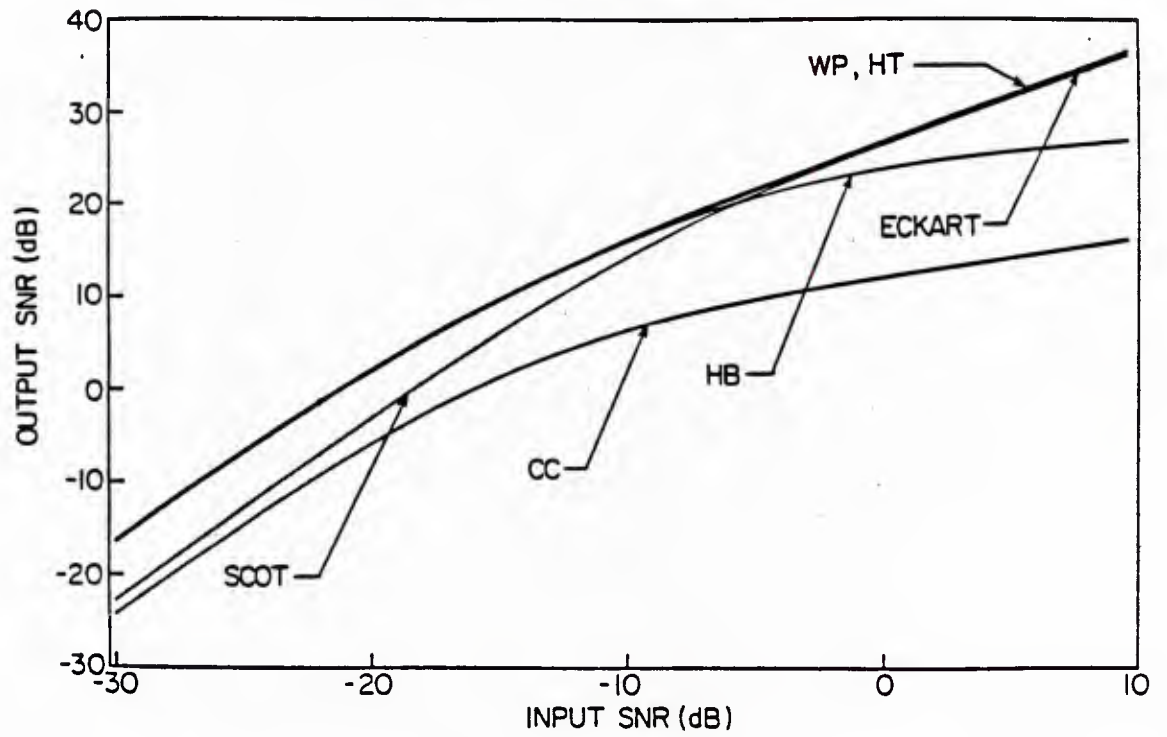
Fig. 6

Performance of various processors under SNR_2
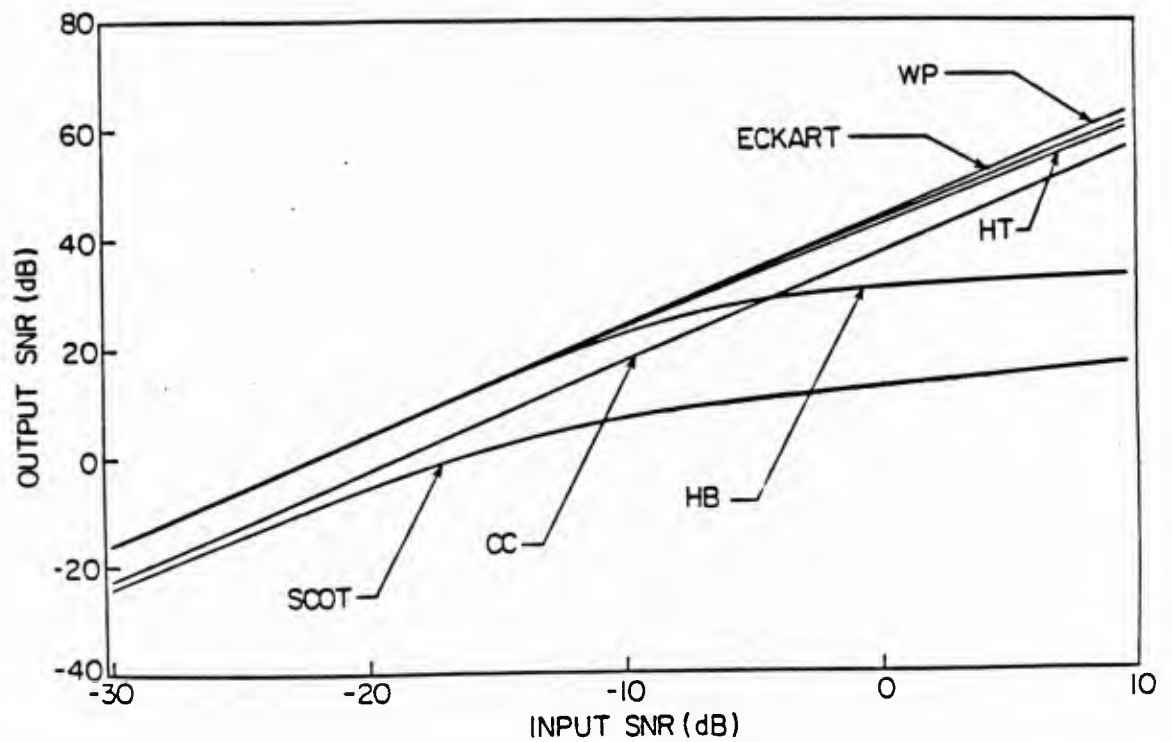
Fig. 7

Performance of various processors under SNR_3
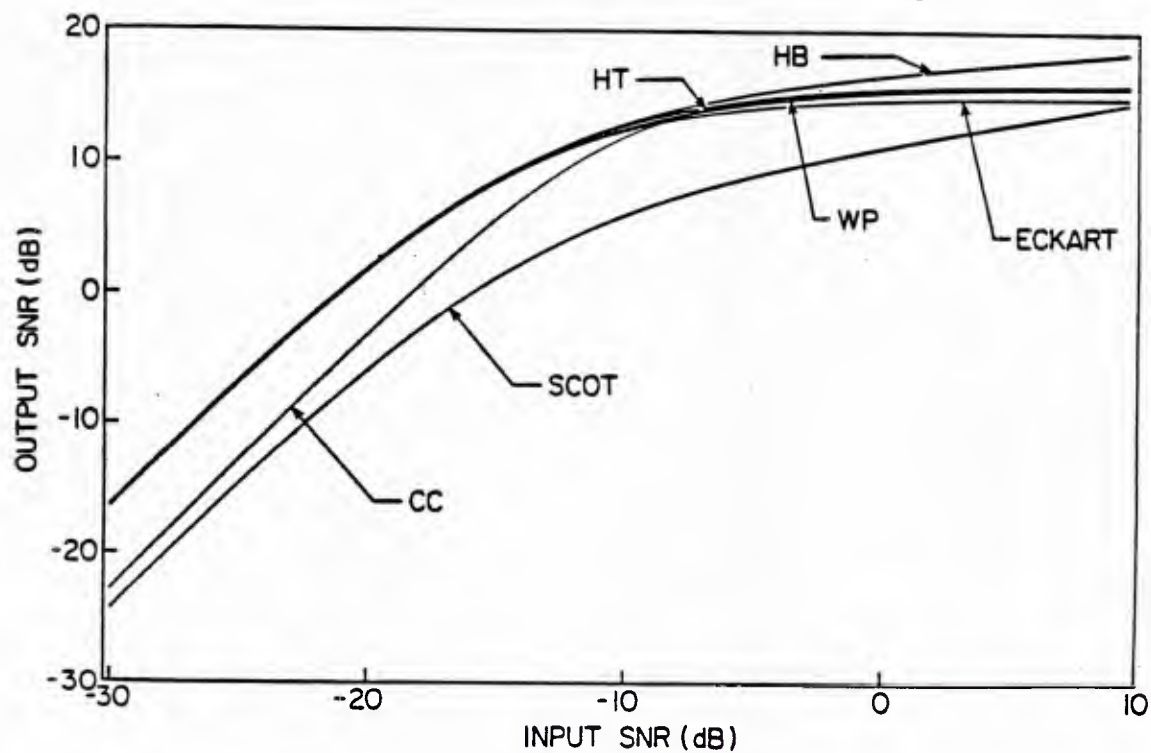
Fig. 8

Local variance for various processors
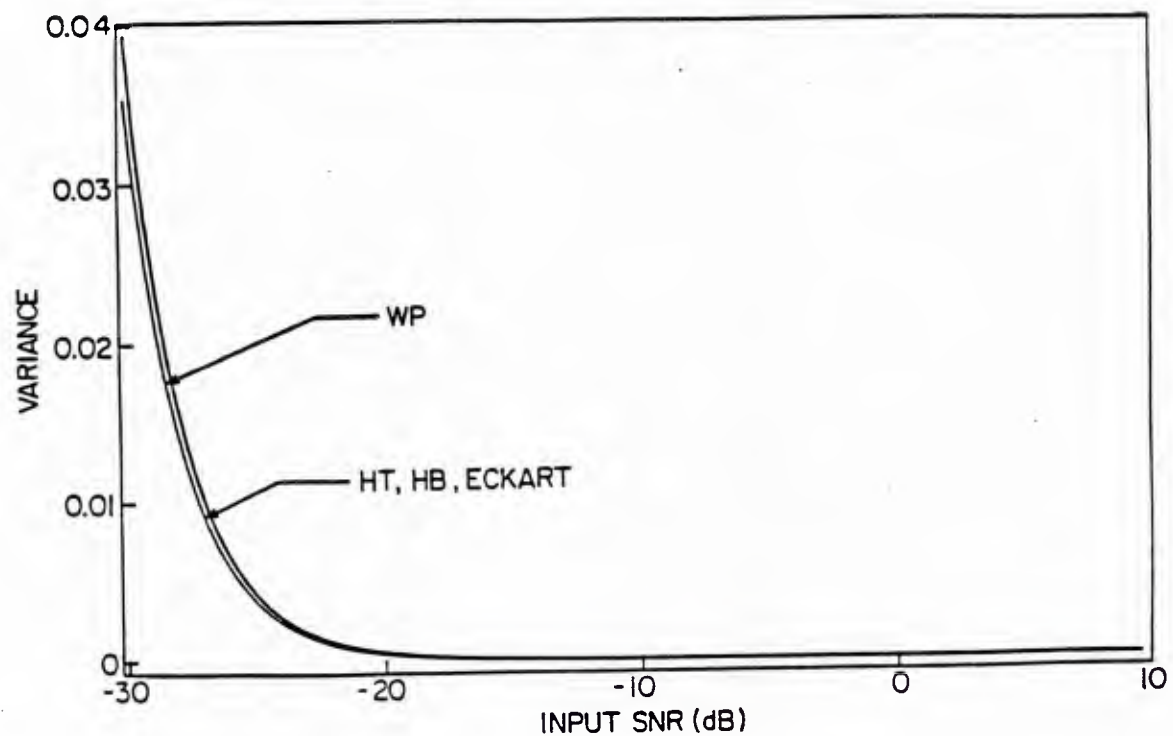
Fig. 9

Nominal signal spectrum for Ex. 2
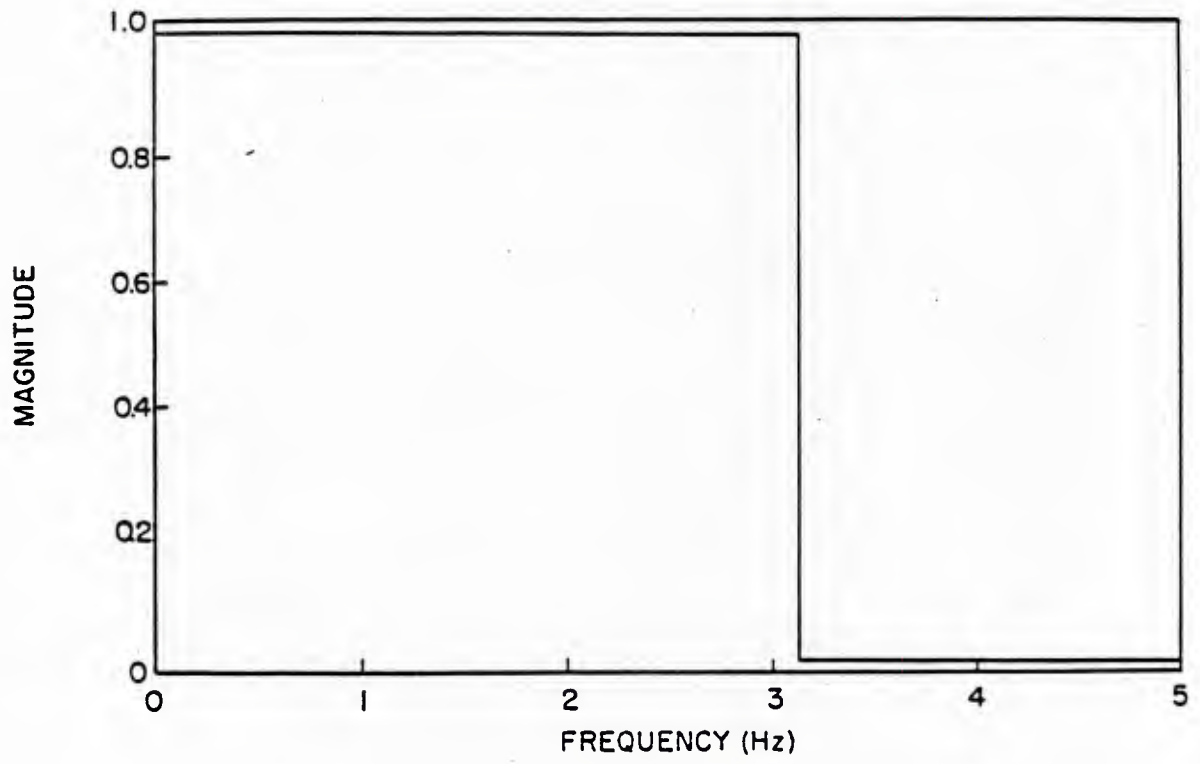
Fig. 10

Nominal noise spectrum for Ex. 2
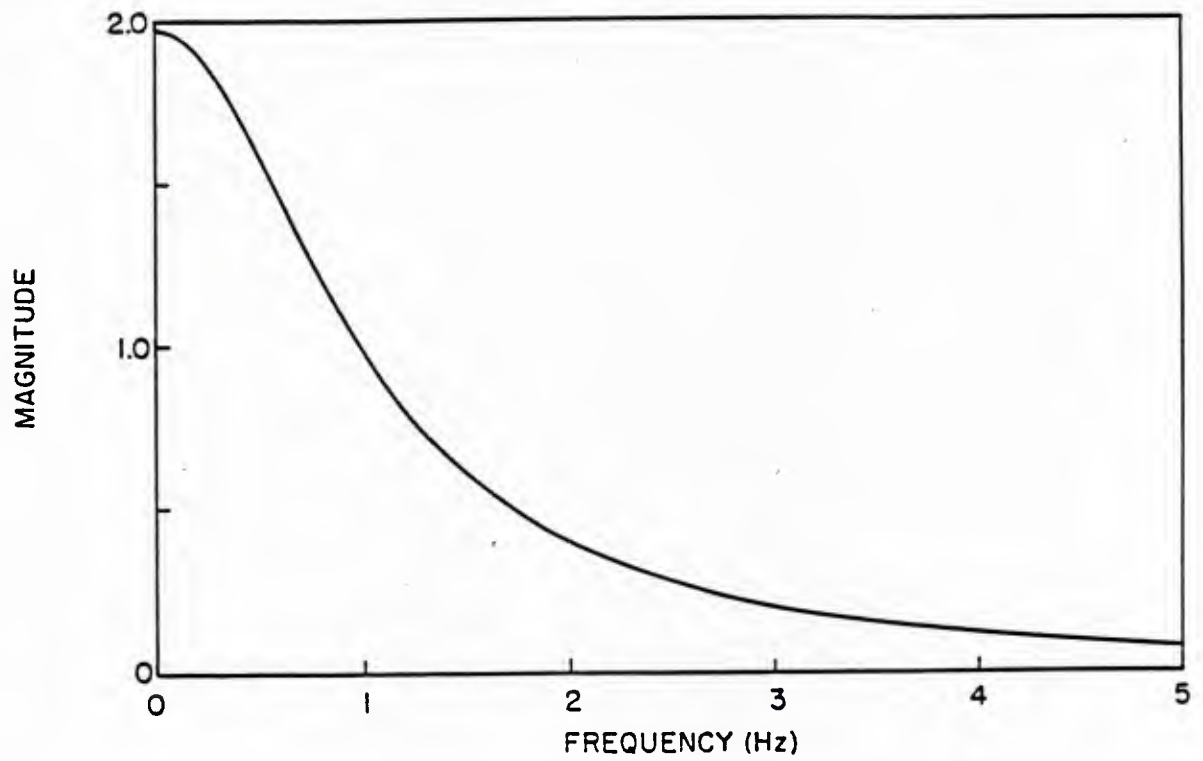
Fig. 11

Least favorable signal spectrum for Ex. 2
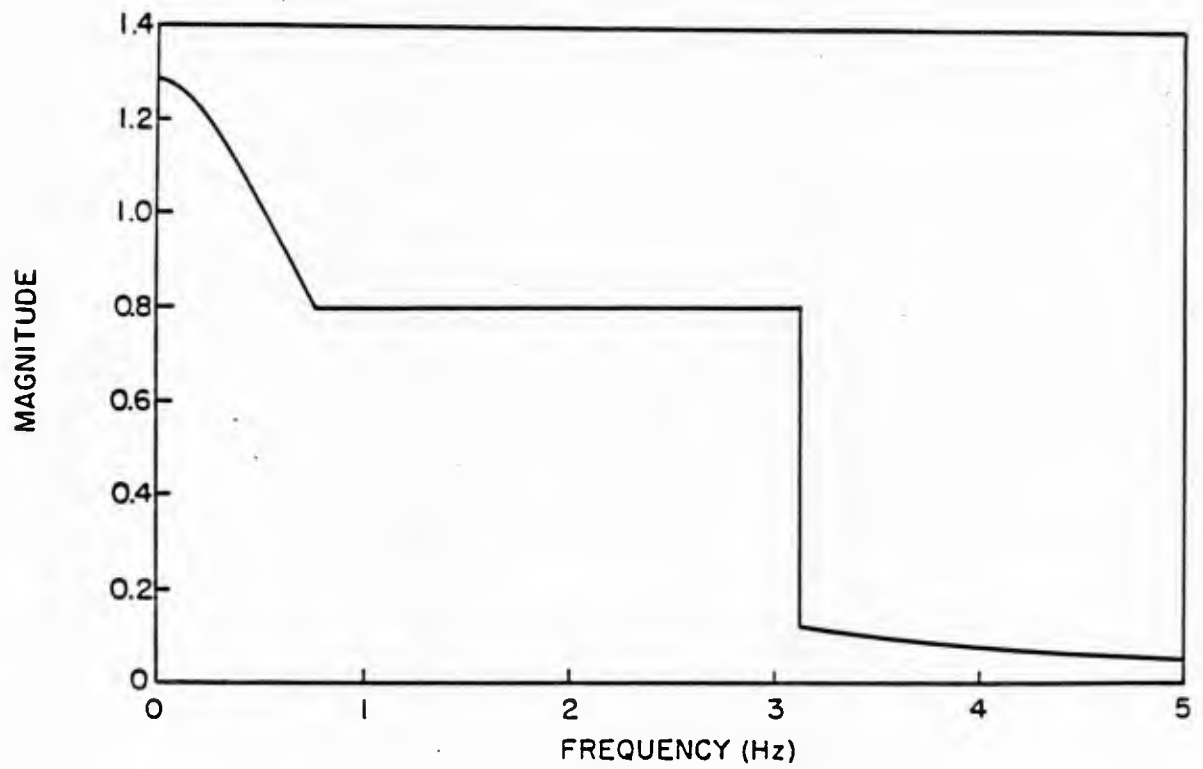
Fig. 12

Least favorable noise spectrum for Ex. 2
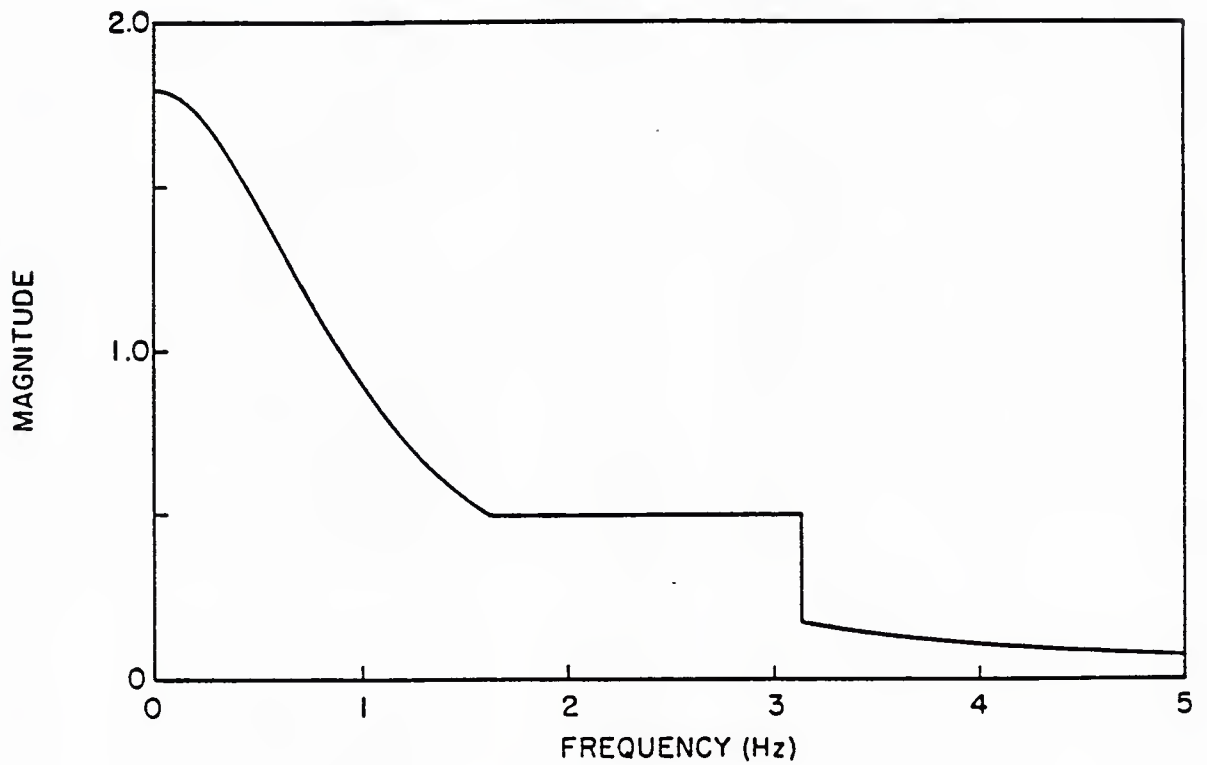
Fig. 13

SNR_1 at the nominal spectra
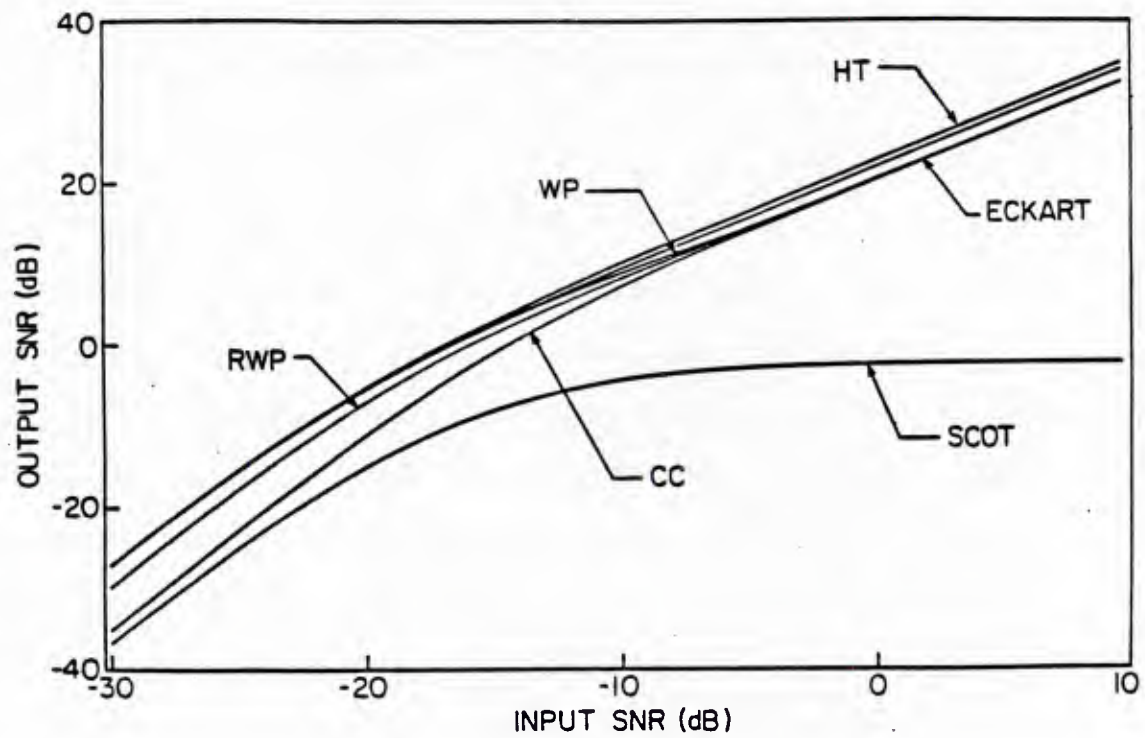
Fig. 14

SNR_1 at least favorable spectra
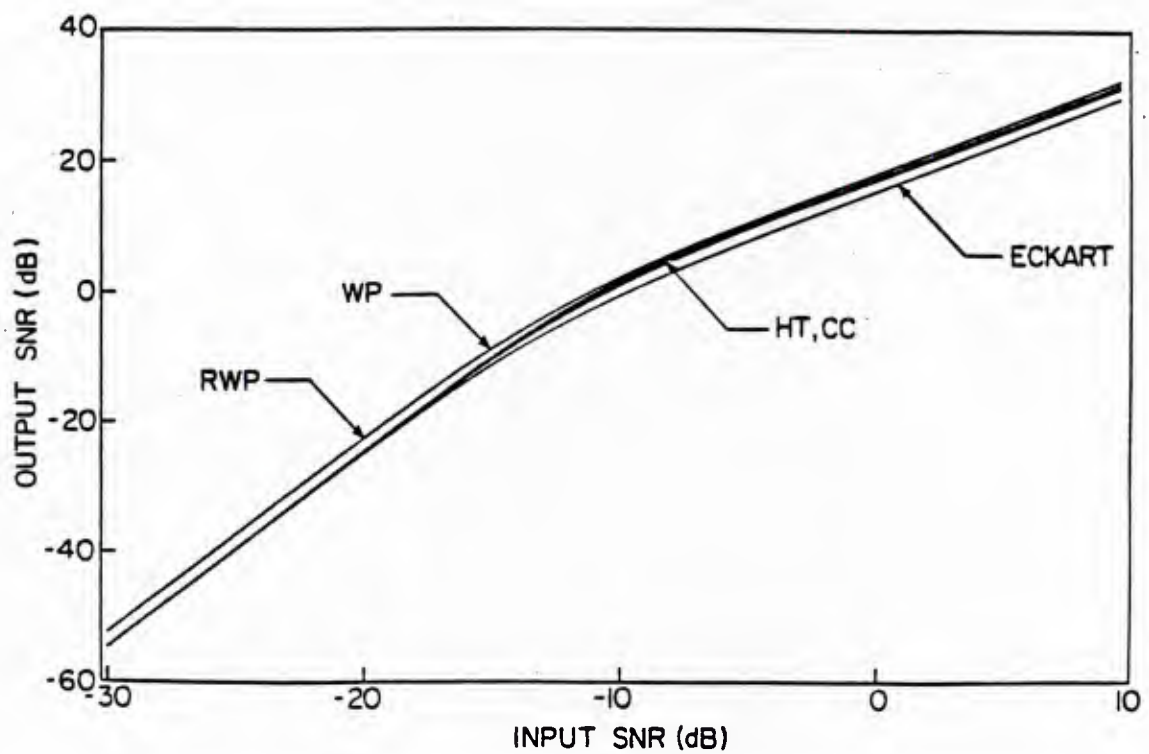
Fig. 15:

Chapter 3
The Center of Symmetry Estimator for Time Delay

I. Introduction

In the previous chapter, the Generalized Cross Correlator (GCC) estimate of time delay, given by the location in time of the global peak of the GCC output trajectory, was motivated from a SNR point of view. In fact, the Maximum Likelihood Estimator (MLE) of the time delay is implementable as a GCC, the Hannan-Thomson processor (HT), for the observation model (1) in Chapter 2 and sufficiently large observation time.. It is well known that, under broad conditions, the MLE asymptotically achieves minimum variance over all unbiased estimators. Hence, in the limit of large observation time, the GCC family, and specifically the HT, has a strong claim to supremacy over other methods. However, for finite observation time, the MLE can only be said to achieve a minimum local variance; that is, when the variation in the estimate is such that it is highly unlikely to fall outside of the immediate vicinity of the true delay at the global maximum. This assumption is especially tenuous when the spectra of the observations contain a narrowband component [2].

In this chapter an alternate estimation scheme is presented which, by design, takes into account non-local error and appears to be less sensitive to narrowband components. The idea is to substitute a center of symmetry estimate (CSE) of the cross-correlation function in place of the global peak estimate used in the conventional GCC. A general procedure is introduced which minimizes the energy in the asymmetric component of the GCC over an *a priori* known region of time delay. This is done by looking at the coefficients in a generalized Fourier expansion of the GCC trajectory. In particular, the Walsh expansion yields simple procedures for estimating symmetry, the Local Center of Symmetry Estimate (LCSE), which operates on the GCC trajectory, and the Modulus CSE (MCSE),

which acts on the absolute magnitude of the GCC waveform. After deriving an approximate expression for the variance of the LCSE, for a general filtered sample cross-correlation function, the filter is optimized to give minimum local variance and is compared to the MLE. The optimal filter takes the form of a simple "matched filter" when the observation time is sufficiently large. The local variance of the MCSE is then derived for the case where the second order distributions of the GCC trajectory are known.

Simulation results for narrowband spectra are next presented which indicate that, for the case considered here, the optimal LCSE has better large error performance, as compared to two popular representatives of the GCC class, the HT and the SCOT.

The chapter is organized as follows. First the problem is stated and the difficulty alluded to above is described when the conventional optimal processor is used. We then present the alternate procedures . A small error approximation to the variance of the LCSE and MCSE estimates is developed analogously to [3]. Optimization of·the CSE weighting is performed for the LCSE based on the local variance approximation and a comparison of the local variance of the optimal GCC versus the LCSE estimator is undertaken. Based on the local variance expression a simple example is presented illustrating the difference in estimation philosophy between the GCC and the LCSE. Finally the results of a simulation of the CSE scheme against the two GCC processors are discussed.

II. Problem Statement

As in Chapter 2, two Gaussian wide sense stationary processes $x_1(t)$ and $x_2(t)$ are observed over a time interval $[0,T]$. However, here it is assumed that the intersensor transfer characteristic is pure delay, so that $|C(w)| = 1$ in (1) of Chapter 2. In other words, $x_1(t)$ and $x_2(t)$ are assumed to be of the form

$$x_1(t) = s(t) + n_1(t) \tag{1}$$
$$x_2(t) = s(t - D_o) + n_2(t)$$

(Note the slight change in notation from Chapter 2 for the delay parameter). Here $s(t)$ is a random signal with an autocorrelation function $R_{ss}(\tau)$ which falls off to zero for τ greater than T_c, the correlation time of the signal. The noises $n_1(t)$ and $n_2(t)$ are uncorrelated and are taken as broadband with respect to $s(t)$. Again assume that the sensor observation time, T, is much greater than both T_c and the maximum *a priori* deviation of D_o, D_m.

The GCC estimate of time delay is, as before, obtained from the global maximum over $[-D_m, D_m]$ of a suitably filtered sample cross-correlation function (Here we assume coherent cross-correlation estimates are used). However, for notational purposes, we will specifically denote the GCC estimate as $\hat{D}_{GCC}$ to distinguish it from the other estimates introduced in this chapter. $\hat{G}_{12}(w)$ is the sample cross-spectrum obtained by Fourier transforming the sample cross-correlation $\hat{R}_{12}(\tau)$. The defining relations for the GCC trajectory and the associated peak estimate of delay are repeated from Chapter 2

$$R^g(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{G}_{12}(w) W(w) e^{jw\tau} dw \qquad \tau \in [-T, T]$$
$$\hat{D}_{GCC} = \underset{\tau \in [-D_m, D_m]}{argmax} \left(R^g(\tau) \right) \tag{2}$$

For the discussion to follow we will need to define the quantity $T_c{}'$ as the smallest number such that for a given α close to zero

$$|R_{ss}^g(\tau)| = \left| \frac{1}{2\pi} \int_{-\infty}^{\infty} |G_{12}(w)| W(w) e^{jw\tau} dw \right| < \alpha \tag{3}$$
$$|\tau| > T_c{}'$$

$T_c{}'$ is interpreted as the α-correlation time of the filtered process

$$s(t) * w^+(t)$$

$$w^+(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} W^+(w)e^{jwt} \, dw$$

and $W^+(w)$ has all of its poles in the LHP.

$$W(w) = W^+(w)\left(W^+(w) \right)^*$$

Under an asumption of sufficiently large observation time, the MLE, for the unknown delay D_o, takes the form of a GCC implemented with the HT weight (Eq. (13) of Chapter 2) [1]. This GCC has also been shown to minimize the local (small error) variance of any GCC estimate of delay [6]. Therefore, the HT will outperform any other estimator for sufficiently large observation time, and will locally outperform any other GCC.

The local variance is a measure which is only sensitive to "small errors"; it characterizes estimator performance for high signal-to-noise ratio in the immediate region of the true delay D_o. However, it significantly underestimates the actual variance when the global peak is likely to be far removed from D_o, as can occur for even moderately high signal-to-noise ratio and narrowband signals using the HT [2]. Thus the optimality of the HT can only be asserted in a small error sense. The maximum value of the trajectory of the HT is in general highly unstable in the sense that small variations in signal-to-noise ratio can translate into large and abrupt changes in the location of the absolute maximum (e.g. waveforms in Fig. 2.1 and 2.2). This discontinuous behavior of the estimate is characteristic of GCC's in general due to the peak detection based estimation procedure. It is this undesirable property of peak detection oriented time delay estimation schemes that motivates the subject of this chapter.

III. Center of Symmetry Estimates

GCC type processors can be interpreted as estimators which utilize the property that, asymptotically, the cross-correlation function assumes its absolute maximum at the true time delay D_o. Here we exploit a different asymptotic property of the cross-correlation to motivate another approach: the delay D_o occurs at the center of symmetry of the cross-correlation function. Essentially the estimation method involves extracting the asymmetric component of R^g about each candidate point $\hat{D}$ in the *a priori* region. Then the center of symmetry estimate is set to that point which minimizes the energy in this asymmetric component. The general procedure for estimating the center of symmetry is stated below.

Let $\{E_k(\tau), O_k(\tau)\}$ be a complete orthonormal set of functions on an interval $[-L, L]$. Here the E_k's are even and the O_k's are odd about the origin. We now expand R^g in a Fourier series over the length $2L$ interval centered at $\hat{D}$, $[\hat{D}-L, \hat{D}+L]$

$$R^g(\tau) = \sum_{i=1}^{\infty} e_i(\hat{D})E_i(\tau-\hat{D}) + \sum_{i=1}^{\infty} o_i(\hat{D})O_i(\tau-\hat{D}) \tag{4}$$

Where we have defined the Fourier coefficients $e_k(\hat{D})$, $o_k(\hat{D})$ of R^g corresponding to $E_k(\tau-\hat{D})$, $O_k(\tau-\hat{D})$

$$e_k(\hat{D}) = \int_{\hat{D}-L}^{\hat{D}+L} R^g(\tau)E_k(\tau-\hat{D})d\tau$$

$$o_k(\hat{D}) = \int_{\hat{D}-L}^{\hat{D}+L} R^g(\tau)O_k(\tau-\hat{D})d\tau$$

(5)

Now the first and second terms to the right of the equality in Eq. (4) represent the even and odd components of R^g about the point $\hat{D}$. Because of the orthonormality of the expansion functions, minimimizing the energy in the odd component over $\hat{D} \in [-D_m, D_m]$ is equivalent to minimizing the sum of the squares of

the $o_i(\hat{D})$. For practical purposes we minimize only the N-th partial sum of the squares of the odd Fourier coefficients to give the center of symmetry estimate

$$\hat{D}_{CSE} = \underset{\hat{D} \in [-D_m, D_m]}{argmin} \sum_{i=1}^{N} |o_i(\hat{D})|^2 \tag{6}$$

As stated above the Center of Symmetry Estimate depends on quite a few free parameters: the set of orthonormal functions $\{E_k, O_k\}$; the "window length" $2L$; the truncation parameter N and the specific GCC weighting employed $W(w)$. In principle one could select the orthonormal functions such that the symmetry information is concentrated in the first few terms of the Fourier series. L should be chosen to minimize the multiplicity of the local minima of the N-th partial sum in Eq. (6) in order to maximize the resolution of the CSE procedure.

Here a specific implementation of the CSE will be analyzed and evaluated. This corresponds to choosing the Walsh functions over the window $[-L, L]$ as the orthonormal set. The first three of these functions are shown in Figs. 3.1 a,b and c. As a particularly simple case we use only the first odd Walsh function in the family, Fig. 3.1b, to form the CSE. That is we take N equal to one in Eq. (6). The procedure then only involves evaluating the simple statistic

$$o_1(\hat{D}) = \frac{1}{2L} \int_{\hat{D}-L}^{\hat{D}+L} R^g(\tau) sgn(\tau - \hat{D}) d\tau \tag{7}$$

Since the normalization by $2L$ has no effect on the location of the minima of o_1 we will often omit it. This particular CSE, which we will call the Local Center of Symmetry Estimate (LCSE), merely takes the center of symmetry as that point, $\hat{D}_{CSE}$, which divides the area under R^g to the left and right of $\hat{D}_{CSE}$ into two equal parts. However due to the truncation of the infinite sum in Eq. (5) to a single term in Eq. (6) the general CSE procedure must be modified to take into account further properties of R^g at its center of symmetry, namely that there is a global maximum there for large SNR. The actual procedure proposed uses the

estimate

$$\hat{D}_{CSE} = \begin{cases} \text{median of the upcrossings} \\ \text{of } o_1 \text{ within } [-D_m, D_m] \end{cases} \tag{8}$$

An upcrossing of o_1 is defined as a point where $o_1(\tau)$ changes sign from negative to positive as τ increases. The median of a monotonic sequence $y_1, y_2, \ldots, y_n$ is defined to be $y_{\frac{n}{2}+1}$ for n even and $y_{\frac{n+1}{2}}$ for n odd.

In regard to Eq. (8) refer to Fig. 3.2. The plausibility of the estimate is made clear by the following. If R^g has a local maximum at its center of symmetry then $o_1(\tau)$ will upcross the origin there since the area identified with "+" becomes greater than that identified with the "-" in Fig. 3.2. In the case where R^g is everywhere non-negative in the integration window and L is such that , asymptotically, $R^g(\tau)$ is concentrated in $[\hat{D}-L, \hat{D}+L]$ for all $\hat{D} \in [-D_m, D_m]$, (e.g. $L \geq T_c' + D_m$), $o_1(\tau)$ gives a unique zero converging to D_o. We are more interested in the case where the left and right tails of R^g do not significantly fall outside of the leftmost and rightmost integration windows $[-D_m-L, -D_m+L]$ and $[D_m-L, D_m+L]$ respectively (See Fig. 3.2). In this case multiple upcrossings occur in o_1. However, asymptotically, they are at locations which are symmetric about the center of symmetry of R^g. Thus the median picks out the correct zero location of Eq. (7) corresponding to the true delay.

In principle the window parameter L can be chosen to optimize the theoretical and/or practical performance of the estimate. Indeed in the limit as L becomes small, $L = \Delta$ say, the LCSE can be looked at as a generalization of a peak discriminator, i.e. a GCC, by noting that, within a scale factor, we have from Eq. (7) (by the mean value theorem)

$$o_1(\tau) = \frac{R^g(\tau+\Delta) - R^g(\tau)}{\Delta} \doteq \frac{dR^g(\tau)}{d\tau} \tag{9}$$

Hence, for asymptotically small L, $o_1(\tau)$ upcrosses zero wherever there is a local peak of R^g, obviously not a good attribute for our purposes. On the other hand if the maximum admissible deviation of D_o from zero, D_m, is such that $D_m + T_c' < T$ then one can integrate over the majority of the positive extent of the normalized R^g by setting $L = T_c'$. This choice should yield a statistic, o_1, which has the least chance of getting hung up on a local maximum of R^g.

In view of the effect of multiple zeros on the LCSE discussed above it would be to our advantage to achieve a reduction in the number of zero upcrossing solutions to Eqns. (6) or (7). This suggests the application of the asymmetric component minimization, described above, to the absolute value of the GCC output trajectory. For definiteness we will concentrate on the LCSE procedure applied to the modulus of R^g. This estimate, denoted the MCSE, is obtained by finding the median of the locations of the zero upcrossings within $[-D_m, D_m]$ of

$$o_1^M(\hat{D}) = \frac{1}{2L} \int_{\hat{D}-L}^{\hat{D}+L} | R^g(\sigma) | \; sgn(\tau - \hat{D}) d\tau \qquad (10)$$

Note that for all of the above CSE methods, unlike the GCC estimates, a small random change in the detailed structure of R^g will not translate into large deviations of $\hat{D}_{CSE}$, since only large changes in the area under $R^g(\tau)$, $\tau > D_o$ or $R^g(\tau)$, $\tau < D_o$ can significantly change the location of a zero of $o_1(\tau)$ or $o_1^M(\tau)$. This can be attributed to the fact that the CSE essentially smooths out spurious peaks in R^g. The next section is concerned with the local or small error properties of the LCSE and MCSE families considered here. Based on these properties the optimal GCC weighting, $W(w)$, for the LCSE, is then derived.

IV. Theoretical Performance of the CSE

Throughout the following we will assume the following for the LCSE and

MCSE and D_m the maximum *a priori* deviation of D_o from zero: 1) $T_c' \ll L \ll T$; and 2) $D_m + L \ll T$. Assumption 1 and 2 essentially constrain the filter $W(w)$ not to overly distort the spectrum $|G_{12}(w)|$ so that the majority of the non-zero range of the mean GCC trajectory always lies within the integration window.

Assuming consistent estimates of $G_{12}(w)$, the GCC trajectory $R^g(\tau)$ will converge to the filtered signal autocorrelation $R_{ss}(\tau) * w(\tau)$, which is symmetric about D_o, as $T \to \infty$. Thus by assumptions 1 and 2, $o_1(\tau)$ and $o_1^M(\tau)$, (Eqns. (7) and (10)) will display zeros symmetrically located about $\tau = D_o$, which implies that the CSE methods are asymptotically locally unbiased.

For comparison purposes, we will derive the local variance of the CSE estimators under analogous small error assumptions used elsewhere for the derivation of the local variance of the GCC [3],[5]. Then the resulting expression will be compared to that of the GCC in the case of the LCSE. Finally an asymptotic approximation to the optimal GCC weight in terms of LCSE variance will be presented.

For simplicity we let $o_1(\tau)$ represent the two CSE statistics for the LCSE, Eq. (7), and the MCSE, Eq. (9). We use essentially the same idea to derive the local variance as in reference [5]. Define the mean value of $o_1(\tau)$, $\bar{o}_1(\tau) = \mathrm{E}[o_1(\sigma)]\,|_{\sigma = \tau}$. Now expand $o_1(\tau)$ in a Taylor series about D_o retaining only the linear term explicitly

$$\bar{o}_1(\tau) = \bar{o}_1(D_o) + (\tau - D_o)\,\frac{\partial \bar{o}_1(D_o)}{\partial D_o} + o(\tau - D_o) \tag{11}$$

Note that $\bar{o}_1(D_o) = 0$.

Assume that $o_1(\tau)$ crosses through zero at $\tau = \hat{D}$ and that $\hat{D}$ lies in the linear region about D_o of $\bar{o}_1(\tau)$, which will be referred to as the local error region, $\mathbf{R_L}$

(See Fig 4.1). That is $\hat{D}-D_o$ is small enough so that $o(\hat{D}-D_o)$ terms in Eqn (11) can be neglected. Take the variance of both sides of Eq. (11) to obtain

$$var[\bar{o}_1(\hat{D})] = var[\hat{D}]\left[\frac{\partial\bar{o}_1(D_o)}{\partial D_o}\right]^2 \qquad (12)$$

Now as a small error approximation, the deviation of $\bar{o}_1(\hat{D})$ for $\hat{D} \in \mathbf{R_L}$ is identical to the deviation of $o_1(\tau)$ at $\tau = D_o$ over the realizations of $o_1(\tau)$ having local zero upcrossings (see Fig. 4.1). Hence, we obtain for the local variance of $\hat{D}$

$$var_L[\hat{D}] = \frac{var[o_1(D_o)]}{\mathbf{E}^2\left\{\frac{\partial o_1(\tau)}{\partial\tau}\right\}|_{\tau=D_o}} \qquad (13)$$

which is analogous to the expression obtained for the local variance of the GCC estimate [5] .

$$var[\hat{D}_{GCC}] = \frac{var\left\{\frac{\partial R^g(D_o)}{\partial D_o}\right\}}{\mathbf{E}^2\left\{\frac{\partial^2 R^g(D_o)}{\partial D_o^2}\right\}} \qquad (14)$$

The denominator of (13) is simply computed

$$\frac{\partial}{\partial\tau}\,o_1(\tau)|_{\tau=D_o} = 2R^g(D_o) - R^g(D_o+L) - R^g(D_o-L) \qquad (15)$$

Since $R^g(t)$ is a locally unbiased estimator of $R_{ss}(\tau) * w(\tau)|_{\tau=t-D_o}$ which is symmetric about D_o

$$\mathbf{E}\left\{\frac{\partial}{\partial\tau}\,o_1(\tau)|_{\tau=D_o}\right\} = [2R_{ss}(\tau) * w(\tau) - 2R_{ss}(\tau-L) * w(\tau-L)]\,|_{\tau=0} \qquad (16)$$

By assumption $L \gg T_c'$ so that the second term in brackets in Eq. (15) is approximately zero. Thus the denominator of Eq. (13) becomes in the frequency domain

$$E\left\{ \frac{\partial}{\partial \tau} \; o_1(\tau) \mid_{\tau = D_o} \right\} = \frac{1}{\pi} \int_{-\infty}^{\infty} \mid G_{12}(w) \mid \; W(w) dw \qquad (17)$$

Even though we wish to apply the CSE procedure to narrowband signals, for the derivation of the local variance it is necessary to assume a large time-bandwidth product. The resulting expression will be seen, however, to well characterize the performance of the CSE even for narrowband signal spectra. Under these assumptions, $R^g(\tau)$ is approximately covariance stationary for $\tau \neq D_o$ (See Appendix C). Define $\sigma(\tau, \nu)$, the covariance function of $R^g(\tau)$ and $R^g(\nu)$, and the GCC spectrum, $C(w)$. $var[o_1(D_o)]$ is computed in Appendix A for the LCSE. For large values of L the variance of the LCSE has the approximate form

$$var_L [\hat{D}] = \frac{L}{2} \; \frac{C(0) \mid W(0)\mid^2}{\left[\frac{1}{\pi} \int_{-\infty}^{\infty} \mid G_{12}(w) \mid \; W(w) \; dw \right]^2} \qquad (18)$$

An expression is derived in Appendix C for $C(w)$, Eqn (C.4), which is, for $w = 0$

$$C(0) = \frac{1}{k} \; \mid W(0)\mid^2 \; (G_{11}(0) G_{22}(0) + \mid G_{12}(0)\mid^2) \qquad (19)$$

where k is directly proportional to T, the observation interval. Substituting (19) into (18) and recalling the definition of the coherency

$$var_L \hat{D} = \frac{L}{2k} \; \frac{G_{11}(0) G_{22}(0)(1 + \mid \gamma_{12}(0)\mid^2) \mid W(0)\mid^2}{\left[\frac{1}{\pi} \int_{-\infty}^{\infty} \mid G_{12}(w) \mid \; W(w) \; dw \right]^2} \qquad (20)$$

Note that since L is a fixed constant in Eq. (20) the LCSE estimate is consistent since $k \to \infty$ as $T \to \infty$.

If the two dimensional distributions of R^g are known then the local variance of the MSCE can be derived (See Appendix B). Define

$$\Psi_t(u) = \mathbf{E}\{e^{iuR'(t)}\}$$
$$\Phi_{t,\sigma}(u) = \mathbf{E}\{e^{iuR'(t)R'(\tau)}\}$$

$$\nu(t,\tau) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{1}{u^2} \Phi_{t,\tau}(u)\,du$$

$$- \left[\frac{1}{\pi} \int_{-\infty}^{\infty} \frac{1}{u^2} \Psi_t(u)\,du \right] \left[\frac{1}{\pi} \int_{-\infty}^{\infty} \frac{1}{u^2} \Psi_\tau(u)\,du \right]$$

The function $\nu(\tau,\sigma)$ is the covariance function between $|R^g(\tau)|$ and $|R^g(\sigma)|$. Substituting Eq. (B.32) in Eq. (13) we obtain for the local variance of the MCSE

$$var_L \hat{D} = \frac{\int_C d\tau \int_C d\sigma\, \nu(\tau,\sigma)}{\left[\frac{1}{\pi} \int_{-\infty}^{\infty} \frac{1}{u^2} \Psi_{D_0}(u)\,du \right]^2} \qquad (21)$$

where $\int_C$ denotes the difference of the two integral operators

$$\int_{D_\bullet-L}^{D_0} - \int_{D_0}^{D_\bullet+L} \qquad (22)$$

Without further assumptions on the distribution of the GCC trajectory, Eq. (21) is too general to be useful in indicating the performance of MCSE.

For comparison we display the expression derived in [5] for the GCC estimator variance

$$var_L \hat{D}_{GCC} = \frac{1}{k} \frac{\frac{1}{2\pi} \int_{-\infty}^{\infty} w^2 G_{11}(w) G_{22}(w) (1 - |\gamma_{12}(w)|^2) |W(w)|^2 dw}{\left[\frac{1}{2\pi} \int_{-\infty}^{\infty} w^2 |G_{12}(w)| W(w)\,dw \right]^2} \qquad (23)$$

The quantity $G_{11}(w)G_{22}(w)(1 - |\gamma_{12}(w)|^2)|W(w)|^2$ can be interpreted as the spectral density of the component of R^g which corrupts the global peak at D_o.

This is the correlator noise in the representation, (3) of Chapter 2, of the GCC output. Comparing the numerators of (20) and (23), it is evident that the variance of the LCSE depends on the correlator noise variance only through the value of its spectrum in the immediate neighborhood of $w = 0$, while for the GCC it varies as the second moment of the noise spectrum. This reflects the time averaging criterion which forms the crux of the center of symmetry estimate. In other words, only the average value of $R^g(\tau)$ to the left and right of $\tau = \hat{D}$ are utilized by $o_1(\tau)$ in its search for a zero. The above comments imply that the (local) variance of the LCSE will be insensitive to changes in the detailed, i.e. high frequency, structure of the noise spectra as long as the average, or D.C., power remains the same. Thus a certain robustness to the underlying noise in the observations is indicated.

Eq. (20) is a limiting form of the actual expression for the LCSE variance derived in Appendix A. In so far as the actual performance is well characterized by this asymptotic expression, it is clear that any weighting function $W(w)$ possessing a null (infinite attenuation) at $w = 0$ is optimal for the LCSE class as long as the denominator of Eq. (20) is finite, i.e. $W(w)$ does not annhilate the signal spectrum $G_{ss}(w)$. In particular a modified simple cross-correlator: $W(w) = 0$ if $G_{ss}(w) = 0$ or if $w = 0$, $W(w) = 1$ otherwise, would satisfy the optimality condition mentioned above.

The above simple optimizing weight arises from consideration of Eq. (20). A more accurate expression for the LCSE variance is given in Appendix A, Eq. (A.11). Using Eq. (A.11) an optimal W, within the class of integrable frequency functions continuous at the origin, $w = 0$, can be derived for large L and the additional assumption of smooth $C(w)$. It has the form

$$W(w) = \frac{G_{ss}(w)}{1 + C(w)\delta_L(w)} \tag{24}$$

where

$$\delta_L(w) = 2L sinc^2(Lw) \tag{25}$$

The filter $W(w)$ of Eq. (24) is derived using the observation that, for sufficiently large L, $\delta_L(w)$ is either less than ϵ or larger than $1/\epsilon$ except over frequency intervals with aggregate length less than $2/\sqrt{\epsilon L}$. In [7, Appendix D] this is shown to imply that the numerator of the local variance expression, (A.11), can be replaced by $k \int\limits_{-\infty}^{\infty} (1+C(w)\delta_L(w))|W(w)|dw$ with only a small error. For L very large, δ_L in Eq. (24) is essentially zero over all frequency except in the neighborhood of the origin. We will investigate the performance of the limiting form of Eq. (24)

$$W(w) = G_{ss}(w) \tag{26}$$

where we have neglected the notch at $w = 0$ of Eq. (24). The filter in (26) is matched to the spectrum of the signal. For convenience, we shall refer to it as a matched filter. This does not define a matched filter in the conventional sense since the signal is non-deterministic and no prewhitening is employed.

The two "optimal" procedures mentioned above will be investigated in the this paper. One, $W(w) \equiv 1$ over the region supporting $G_{ss}(w)$, will be referred to as the cross-correlator (CC) and the other, given by Eq. (26), the matched filter (MF) versions of the LCSE, or the CC-LCSE and MF-LCSE respectively. We now present a simple example for the optimal LCSE in the case where the observations have flat spectra.

Assume the signal and noise spectra are flat and strictly bandlimited

$$
\begin{aligned}
G_{ss}(w) &= S \frac{2\pi}{B} && |w| \le B/2 \\
&= 0 && |w| > B/2 \\
G_{ni}(w) &= N \frac{2\pi}{B_n} && |w| \le B_n/2 \\
&= 0 && |w| > B_n/2
\end{aligned}
\tag{27}
$$

For the above case the CC-LCSE and the MF-LCSE give identical local variance expressions. An optimal weight is thus $W(w) = 1$ within the non-zero region of $G_{ss}(w)$ and $W(w) = 0$ elsewhere. From Eq. (18) we have the optimal LCSE local variance

$$var_L \hat{D} = \frac{1 + |\gamma_{12}|^2}{|\gamma_{12}|^2} \frac{\pi^2 L}{2kB^2} \tag{28}$$

and for the H.T. processor, the optimal local variance for peak detection is

$$var_L D_{H.T.} = \frac{1 - |\gamma_{12}^2|}{|\gamma_{12}|^2} \frac{24\pi}{kB^3} \tag{29}$$

The local variance of the LCSE is in general higher than that for the H.T. processor since the H.T. achieves the Cramer-Rao Lower Bound locally. As expected from the preceeding comments, the presence of a high signal-to-noise ratio, i.e. high coherency, will favor the H.T. because of the term $1 - |\gamma_{12}(w)|^2$ in (29). However, at low signal-to-noise ratio, both the optimal LCSE and the H.T. variance behave as the inverse of the coherency. It is important to note that the expression for the H.T. variance behaves as inverse bandwidth cubed while for the LCSE it behaves as inverse bandwidth squared. Consequently, as far as the small error performance is concerned, the LCSE should perform comparably to the H.T. in cases where the signal-to-noise ratio is moderate and the bandwidth of the signal is small. Because of the instability of the H.T. estimate for precisely this combination of signal-to-noise ratio and bandwidth (see the discussion in Section II), the expression for the estimator variance, Eq. (29), can underestimate the true variance of the estimate by orders of magnitude [2],[4]. Due to the smoothing effect of integrating the GCC trajectory, the local variance of the CSE, Eq. (28), may more accurately reflect the actual performance of the estimator than does Eq. (29) for the GCC. Therefore the potential for superior performance of the LCSE scheme may well be underestimated by Eqs. (28) and (29).

V. Simulation Results

For the present we are interested in the performance of the LCSE scheme for narrowband signal spectra, that is the low time-bandwidth product situation. A preliminary experimental investigation of the relative performance of the LCSE and the GCC is reported here.

We synthesized two sensor sequences which contain a common narrowband component in two uncorrelated white Gaussian noises to exercise and compare the HT and the SCOT processors with the CC-LCSE and MF-LCSE discussed in Sections III and IV. Specifically we first generated a unit power Gaussian signal $r(k)$ with periodic spectrum on $[-\pi, \pi]$ via the autoregression

$$r(k) = h_1 \, r(k\text{-}8) + h_2 \, r(k\text{-}16) + a \, w(k) \ , \qquad k = 1, 2, \ldots \qquad (30)$$

where $w(k)$ is a white Gaussian pseudonoise sequence and $h_1 = 1.77$ $h_2 = -0.86$ and $a = 0.17$. Then $r(k)$ was passed through a sixth order elliptic filter to yield a lowpass signal $z(k)$. Finally $z(k)$ was modulated onto a sinusoid to give the signals $s(k)$ and $s(k\text{-}D_o)$

$$s(k) = \sqrt{2}(s_d + s_r \, z(k))sin(2\pi f k + \phi) \qquad k = 1, 2, \ldots, 5120 \qquad (31)$$

Here $f = 0.017$, ϕ is a random initial phase uniform on $[-\pi,\pi]$ and s_d and s_r are the power in the deterministic carrier component and the random component of $s(k)$ respectively. s_d and s_r were chosen to give a carrier-to-signal power ratio (CSR) of -20dB for the amplitude modulated (AM) waveform in Eq. (31). The spectrum and autocorrelation function of $s(k)$ are shown in Figs. 5.1. and 5.2 respectively. $s(k)$ and $s(k\text{-}D)$ were then contaminated with mutually uncorrelated zero mean white Gaussian noises $n_1(k)$ and $n_2(k)$ to form the two records

$$x_1(k) = s(k) + n_p \, n_1(k) \qquad (32)$$

$$k = 1, 2, \ldots 5120$$

$$x_2(k) = s(k\text{-}D) + n_p \, n_2(k)$$

D_0 was set to 25 for the duration of the simulation. The 5120 samples contained in the above two records were divided into 5 distinct groups and the averaged cross-spectral estimate was constructed using the Bartlett procedure. This estimate was then weighted with the HT, SCOT and MF weights displayed in Eqns. (5), (6) and (24). Since the autocorrelation function of the signal $s(k)$ does not fall off to zero within [-511, 512], see Fig. 5.2, the local center of symmetry estimator was implemented. The maximum length integration window was used for the LCSE when *a priori* the true delay lies within 50 bins of the 0-th bin. That is we summed over 451 bins to the left and to the right of each point in the interval [-50,50]. For the SCOT and HT, peak detection was performed and for the MF-LCSE and CC-LCSE, Eq. (7) was implemented in discrete time over the *a priori* region. The above procedure was repeated 100 times for each of 25 signal-to-noise ratios (SNR) between -50 and 40 dB.

Three statistics were computed at each SNR for each processor: the square root of the mean-square-error (MSE) of the estimates; the mean; and the 8 bin frequency deviation, i.e. the number of estimates falling within 8 units of the true delay at 25. These curves are shown for the three schemes: the SCOT; the HT; and the MF-LCSE in Figs. 5.3 through 5.5 respectively.

Referring to Fig. 5.3, for SNR above about -5dB; large errors become insignificant and the HT outperforms the others; but by only 1 bin in the case of the MF-LCSE. As the SNR falls below the small error threshold the MSE performance of the HT deteriorates rapidly with respect to that of the MF-LCSE. The curves in Fig. 5.4 indicate that, on the average, below the -5dB threshold the SCOT has already lost all delay information, giving randomly distributed estimates across the entire interval [-50,50]. Indeed the SCOT procedure breaks down much sooner than either the H.T. or the MF-LCSE, exhibiting a threshold at least 10dB higher than the others. The MF-LCSE picks up a greater amount

of the underlying structure in the GCC waveform which prevents the type of deterioration evident in the case of the SCOT. The performance of the HT, relative to the average of the estimates, falls somewhere in between the SCOT and MF-LCSE. Finally Fig. 5.5 illustrates the performance of the three schemes relative to the small error performance measure, the frequency of the less than 8 bin deviations. For each of the 100 experiments, we plot the percentage of estimates within 8 bins on either side of the true, bin 25. This gives an indication of the percentage of time that the estimate can be expected to give only small errors. Except at the threshold, where a small increase in performance of the MF-LCSE is evident, the HT and the MF-LCSE have virtually identical small error performance, both significantly outperforming the SCOT at all SNRs tested.

In Figs. 5.6 and 5.7 the CC-LCSE experimental performance is displayed. From Fig. 5.6 it is evident that there is virtually no gain in MSE performance over the HT; however it does outperform the SCOT at high SNR. The mean curves, Fig. 5.7, indicate that at low SNR, the CC-LCSE still takes advantage of symmetry structure of the GCC trajectory to yield, on the average, better estimates (by 2 or 3 bins) than the GCCs. The poor performance of the CC-LCSE, and the other GCC's investigated here, reflects the sub-optimality of any $W(w)$ which does not specifically attempt to maximize the denominator of Eq. (20).

The modulus CSEs (MCSE) were also included in the simulation. However they failed to give solutions to the zero location problem, Eq. (10), for about half of the 100 tests at low SNRs and the results are omitted from this study. Of course this is not a conclusive indication of the potential performance of the modulus CSE techniques since the MF and CC weights are not optimal in the sense of minimizing local variance. The specific optimization of the MCSE variance expression should yield estimators which give a more favorable standing to the MCSE than that resulting from the present simulation study.

VI. Discussion

In this chapter several forms of the center of symmetry estimator were presented, which are simple modifications of the Generalized Cross Correlation method for time delay estimation. A particularly simple family of CSE procedures, the LCSE's, was obtained by using a Walsh expansion to characterize the asymmetric part of the GCC trajectory. Based on an optimization of the variance of the LCSE over GCC filters, $W(w)$, an optimal filter function, the "matched filter," was derived. Even though the target performance functional, the local variance of the CSE, may not be valid for large errors, the resultant optimization produced estimates which seem to possess better large error performance than the GCC techniques for narrowband signal spectra. It is believed that this follows from the fact that the LCSE is less sensitive to peak ambiguities than the GCC for narrowband spectra.

The local variance analysis indicates that in general the GCC is a more accurate estimator for small errors than the CSE. This is also evident for the LCSE from the simulation results discussed in the preceeding section. One possible estimation strategem would be to use the GCC and the CSE in a hybrid manner. Using the CSE, one could obtain initial rough estimates for the location of the (local) peak in the GCC trajectory which actually corresponds to the true time delay. This would be a variant of a "gated mode" implementation of the GCC. That is one would censor the GCC output outside of some region determined by the rough estimate obtained using the CSE, and estimate the delay as the location of the highest peak within the gated region. Further study of this and other uses of the CSE remains to be undertaken.

APPENDICES

Chapter 3

Appendix A

Here the expression Eq. (18) is computed assuming $R^g(\tau)$ has covariance given by (C.13) in Appendix C. First compute the variance of $o_1(D_o)$ defined in Eq. (7).

$$var[o_1(D_o)] = var\left[\int_{D_o-L}^{D_o} R^g(\tau)d\tau - \int_{D_o}^{D_o+L} R^g(\tau)d\tau\right]$$

(A.1)

$$= \left[\int_{D_o-L}^{D_o}\int_{D_o-L}^{D_o} + \int_{D_o}^{D_o+L}\int_{D_o}^{D_o+L} - \int_{D_o-L}^{D_o}\int_{D_o}^{D_o+L} - \int_{D_o}^{D_o+L}\int_{D_o}^{D_o}\right] d\tau d\sigma \ \ cov[R^g(\tau),R^g(\sigma)]$$

In Appendix C it is shown that for sufficiently broadband signals, $s(t)$, R_{12} is covariance stationary except at points τ very near D_o, say $\tau \in [D_o^-, D_o^+]$. In any case the difference between the covariance functions involving R^g within the immediate neighborhood of the true time delay and outside of that neighborhood is bounded by a factor of two $(cov[R_{12}(D_o),R_{12}(D_o+\tau)]/cov[R_{12}(\eta),R_{12}(\eta+\tau)] \le 2$, $\eta \ne D_o)$. Hence we can consider the evaluation of the integrals in Eq. (A.1) over regions which do not include $[D_o^-,D_o^+]$ without significant change in the resulting variance expression if the signal is broadband. Accordingly define the spectrum $C(w)$ of $R_{12}(\tau)$

$$C(w) = \int_{-\infty}^{\infty} cov[R_{12}(\eta),R_{12}(\eta+\tau)] \ e^{-jw\tau}d\tau$$

R^g then has the spectrum $C(w)|W(w)|^2$ and the correlation function

$$\rho(\tau) = cov[R^g(\sigma),R^g(\sigma+\tau)] = \frac{1}{2\pi} \int_{-\infty}^{\infty} C(w)|W(w)|^2 dw$$

We can rewrite Eq. (A.1) with the obvious definition of regions γ_1, γ_2, γ_3 and γ_4

to obtain

$$var[o_1(D_o)] = \left\{ \int_{\gamma_1} + \int_{\gamma_2} - \int_{\gamma_3} - \int_{\gamma_4} \right\} dA \ \rho(\tau-\sigma) \qquad (A.2)$$

where $dA = d\tau \, d\sigma$.

Making the change of variable $v = \tau-\sigma$, the regions γ_1, γ_2, γ_3 and γ_4, in Fig. A.1, are mapped to the regions γ_1', γ_2', γ_3' and γ_4', in Fig. A.2, where, for purposes of clarity, D_o is assumed to be positive. We thus have

$$var[o_1(D_o)] = \left\{ \int_{\gamma_1'} + \int_{\gamma_2'} - \int_{\gamma_3'} - \int_{\gamma_4'} \right\} dA' \cdot \rho(v) \qquad (A.3)$$

where $dA' = d\tau \, dv$.

We will deal with the four above integrals individually by integrating first over τ.

Consider the integral over the region γ_1' in Eq. (A.3)

$$\int_{\gamma_1} \rho(v)dA = \int_{-L}^{0} dv \int_{D_o-L}^{\tau+D_o} d\tau \, \rho(v) + \int_{0}^{L} dv \int_{\tau+D_o-L}^{D_o} d\tau \, \rho(v)$$

$$= \int_{-L}^{L} (L - |v|) \ \rho(v)dv \qquad (A.4)$$

Likewise for the region γ_2'

$$\int_{\gamma_2} \rho(v)dA = \int_{-L}^{0} dv \int_{D_o}^{\tau+D_o+L} d\tau \, \rho(v) + \int_{0}^{L} dv \int_{\tau+D_o}^{D_o+L} d\tau \, \rho(v)$$

$$= \int_{-L}^{L} (L - |v|) \ \rho(v)dv \qquad (A.5)$$

And for γ_3'

$$\int_{\gamma_3} \rho(v)dA'$$

$$= \int_{-2L}^{-L} dv \int_{D_o-L}^{\tau+D_o+L} d\tau \, \rho(v) + \int_{-L}^{0} dv \int_{\tau+D_o}^{D_o} d\tau \, \rho(v)$$

$$= \int_{-2L}^{-L} (v+2L)\, \rho(v)\ dv + \int_{-L}^{0} (-v)\, \rho(v)\ dv \tag{A.6}$$

In a similar manner for the region γ_4'

$$\int_{\gamma_4'} \rho(v)\,dA'$$

$$= \int_{L}^{2L} (-v + 2L)\, \rho(v)\ dv + \int_{0}^{L} v\, \rho(v)\ dv \tag{A.7}$$

Note that $\rho(v) = 0$ for $\mid v \mid > T_c'$, where T_c' is defined in Eq. (3). Therefore if $L \gg T_c'$ the first term in Eqns. (A.6) and (A.7) are negligible. Performing the indicated algebra in Eq. (A.3) using Eqns. (A.4) - (A.7) with the above assumption we obtain

$$var[o_1(D_o)] = 2L \int_{-L}^{L} \left(1 - \frac{\mid v \mid}{2L}\right) \rho(v)\ dv \tag{A.8}$$

or

$$var[o_1(D_o)] \approx 2L \int_{-L}^{L} \rho(v)\ dv \tag{A.9}$$

Finally, combining Eq. (A.9) with Eq. (17) and substituting into Eq. (13), we have

$$var_L[\hat{D}] = \frac{L}{2} \frac{\int_{-L}^{L} \left(1 - \frac{\mid v \mid}{2L}\right) \rho(v)\ dv}{\left(R_{ss}(\tau) * w(\tau)\mid_{\tau=0}\right)^2} \tag{A.10}$$

Since ρ is symmetric and $L \gg T_c'$ the numerator of Eq. (A.10) can be considered as the convolution of $\rho(\tau)$ and $(1 - |\tau|/L)$ evaluated at $\tau = 0$. Or in the frequency domain we have

$$var_L[\hat{D}] = \frac{L}{2} \frac{\frac{1}{2\pi} \int_{-\infty}^{\infty} \delta_L(w) C(w) \mid W(w) \mid^2 dw}{\left[\frac{1}{2\pi} \int_{-\infty}^{\infty} G_{ss} W(w)\, dw\right]^2} \tag{A.11}$$

where

$$\delta_L = 2L sinc^2(wL) \tag{A.12}$$

$$sinc(x) = \frac{\sin(x)}{x} \tag{A.13}$$

In the limit of large values of L Eq. (A.10)

$$var_L[\hat{D}] = \frac{L}{2} \frac{C(0)|W(0)|^2}{\left[\dfrac{1}{2\pi} \displaystyle\int_{-\infty}^{\infty} G_{ss} W(w) dw\right]^2} \tag{A.14}$$

Appendix B

Here the local variance of the MCSE, Eq. (21), is derived when the second order distributions of R^g are known. We first need a result on the characteristic function of the modulus of a random variable.

Let X be a random variable with density function $f_X(x)$ and infinitely differentiable characteristic function $\phi_X(t)$. Then the characteristic function, $\phi_{|X|}$, of the absolute value of X is given by

$$\phi_{|X|}(t) = \text{Re } \phi_X(t) + i\, H\left(\text{Re } \phi_X(t)\right) \tag{B.1}$$

where $H\left(\text{Re } \phi_X(t)\right)$ is the Hilbert transform of the real part of $\phi_X(t)$

$$H(z(t)) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{z(u)}{t-u}\, du \tag{B.2}$$

Furthermore, the moments of $|X|$ are given by

$$E\{|X|^{2n}\} = \frac{\partial^{2n} \phi_X(t)}{\partial t^{2n}}(-1)^n \tag{B.3}$$

$$E\{|X|^{2n-1}\} = -(-1)^n \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{(2n-1)!}{u^{2n}}\, \phi_X(u) du$$

To show this we first need a couple of preliminary results. One form of the Dirac

delta function is the integral (Cauchy principal value)

$$\int_{-\infty}^{\infty} e^{itz} dz = 2\pi\delta(t) \tag{B.4}$$

which has the property that for any integrable function $g(t_o)$

$$\int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} e^{i(t-t_0)z} dz \right] g(t)dt = g(t_0) \bullet 2\pi \tag{B.5}$$

We can obtain a useful characterization of Eqn (B.4) in the context of Cauchy integration in the following way.

Consider the contour integral for $t < 0$, $\int_C e^{itz} dz$, where z is a complex variable and C is the contour of Fig. B.1.

Fig. B.1
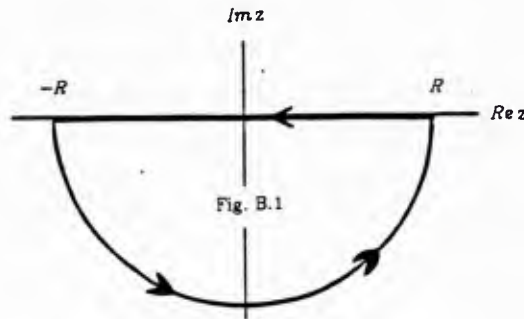
Then by the residue theorem

$$\frac{1}{2\pi i} \int_C e^{itz} dz = 0 \tag{B.6}$$

$$= \frac{1}{2\pi i} \int_{-R}^{R} e^{itz} dz + \frac{1}{2\pi i} (iR) \int_{\pi}^{2\pi} e^{i(tRe^{i\theta}+\theta)} dv$$

Now

$$| e^{i(tRe^{i\theta}+\theta)} | \leq | e^{i\theta+itR \cos\theta} | e^{-tR \sin\theta} \tag{B.7}$$

$$\leq e^{tR} \quad for \ \theta \in [\pi, 2\pi]$$

Therefore combining Eq. (B.6) and inequality (B.7) and taking the limit

$$\left|\lim_{R\to\infty}\int_{-R}^{R}e^{itx}\,dx\right|\leq\lim_{R\to\infty}Re^{tR}=0 \tag{B.8}$$

for $t < 0$.

For $t > 0$, just use the mirror image of the contour in Fig. B.1 to obtain the identical result.

For the case $t \equiv 0$

$$\lim_{R\to\infty}\int_{-R}^{R}e^{itx}\,dx=\lim_{R\to\infty}2R \tag{B.9}$$

Or comparing Eqns (B.4) and (B.9), we have the identity

$$\lim_{R\to\infty}1(t)\,R=\pi\delta(t) \tag{B.10}$$

where

$$1(t)=\begin{cases}1 & t=0\\0 & t\neq0\end{cases} \tag{B.11}$$

The above identity will enable us to show the following fact

$$\int_{0}^{\infty}e^{iux}\,dx=\pi\delta(u)+i\,\frac{1}{u}\,\bigl(1-1(u)\bigr) \tag{B.12}$$

where the imaginary part of Eq. (B.12) has the interpretation: $\frac{1}{u}\bigl(1-1(u)\bigr)$ is equal to zero for $u=0$ and equal to $\frac{1}{u}$, otherwise. Consider the function of a complex variable: e^{iuz}. This has a pole at $z=-sgn(u)\,i\infty$.

For $u > 0$, we consider the contour integral over the path C shown in Fig. B.2.

Fig. B.2
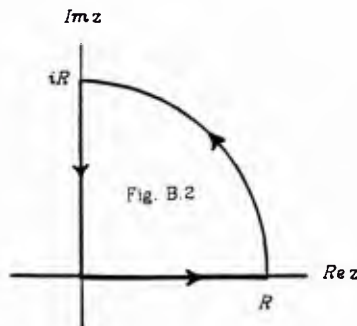
Then we have

$$\frac{1}{2\pi i} \int_C e^{iuz}\, dz = 0 \tag{B.13}$$

$$= \frac{1}{2\pi i} \int_0^R e^{iux}\, dx + \frac{1}{2\pi i} \int_c e^{iuz}\, dz - \frac{1}{2\pi i} \int_0^{iR} e^{iux}\, dx$$

$$= I_R + I_1 + I_2 \tag{B.14}$$

First compute I_1

$$I_1 = \frac{1}{2\pi i} \int_0^{\pi/2} e^{iuRe^{i\theta}} Re^{i\theta}\, id\theta \tag{B.15}$$

$$= \frac{Ri}{2\pi i} \int_0^{\pi/2} e^{-uR \sin\theta}\, e^{i(uR \cos\theta + \theta)} d\theta$$

Note, that since $u > 0$, $I_1 \to 0$ as $R \to \infty$. I_2 is however, non-zero

$$I_2 = \frac{-1}{2\pi i} \frac{1}{iu} \left[e^{-uR} - 1\right] \tag{B.16}$$

or as $R \to \infty$

$$I_2 \to \frac{i}{u} \frac{1}{2\pi i} \tag{B.17}$$

giving the result from Eq. (B.14)

$$\tag{B.18}$$

$$\lim_{R \to \infty} \frac{1}{2\pi i} \int_C e^{iuz}\, dz = \frac{1}{2\pi i} \int_0^\infty e^{iux}\, dx - \lim_{R \to \infty} \frac{R}{2\pi i}\, 1(u) - \frac{1}{2\pi i}\, i\, \frac{1}{u}\, \left(1 - 1(u)\right)$$

We use Eqns. (B.13) and (B.10) to obtain from (B.18)

$$\int_0^\infty e^{iux}\, dx = \pi\delta(u) + i\, \frac{1}{u}\, \left(1 - 1(u)\right) \tag{B.19}$$

Note that formally

$$\int_{-\infty}^\infty e^{iux}\, dx = \int_0^\infty e^{iux}\, dx + \left(\int_0^\infty e^{iux}\, dx\right)^* = 2\pi\delta(u) \tag{B.20}$$

as expected.

The result for $u < 0$ falls out of the above by considering

$$\int_0^\infty e^{-i|u|x}\,dx = \left[\int_0^\infty e^{i|u|x}\,dx\right]^* = \pi\delta(u) - i\frac{1}{|u|}\;\big(1-1(u)\big) \qquad (B.21)$$

$$= \pi\delta(u) + i\,\frac{1}{u}\;\big(1-1(u)\big)$$

Now we show the main result, (B.1) and (B.3), for the characteristic function of the modulus of a random process X with $f_X(x)$ and $\phi_X(t)$ the density and characteristic function of X, that is, $X \sim f_X(x)$, and $\phi_X(t) = \mathbf{E}\{e^{itX}\}$.

$$\phi_{|X|}(t) = \int_0^\infty e^{ixt} f_X(x)dx + \int_{-\infty}^0 e^{-ixt} f_X(x)dx \qquad (B.22)$$

$$= \int_0^\infty e^{ixt}\,[\,f_X(x) + f_X(-x)\,]\,dx$$

$$= \int_0^\infty e^{ixt}\left\{\frac{1}{2\pi}\int_{-\infty}^\infty e^{-ixu}\,[\,\phi_X(u) + \phi_X^*(u)\,]\,du\right\}dx$$

$$= \frac{1}{\pi}\int_{-\infty}^\infty \mathrm{Re}\,\phi_X(u)\left\{\int_0^\infty e^{ix(t-u)}dx\right\}du$$

By (B.12), for $t - u \neq 0$;

$$\int_0^\infty e^{ix(t-u)}dx = \pi\delta(t-u) + i\,\frac{1}{t-u} \qquad (B.23)$$

Hence,

$$\phi_{|X|}(t) = \mathrm{Re}\{\phi_X(t)\} + i\,\frac{1}{\pi}\int_{-\infty}^\infty \frac{\mathrm{Re}\big(\phi_X(u)\big)}{t-u}\,du \qquad (B.24)$$

$$= \mathrm{Re}\{\phi_X(t)\} + i\,H\big(\mathrm{Re}\{\phi_X(t)\}\big)$$

It is well known that if X has moments of all orders

$$\frac{\partial^n}{\partial t^n} \; \mathrm{E}\{e^{iXt}\} \mid_{t=0} = (i)^n \; \mathrm{E}\{X^n\} = \begin{cases} (-1)^{\frac{n}{2}} \; \mathrm{E}\{X^n\} & n \; even \\ (-1)^{\frac{n-1}{2}} \; i \, \mathrm{E}\{X^n\} & n \; odd \end{cases}$$

Therefore, we can calculate the moments of $|X|$

$$(-1)^n \; E\{|X|^{2n}\} = \frac{\partial^{2n} \mathrm{Re}\{\phi_X(t)\}}{\partial t^{2n}} \mid_{t=0} + i \int_{-\infty}^{\infty} \frac{(2n)!}{(-u)^{2n+1}} \; \mathrm{Re}\phi_X(u) du \quad (B.25)$$

By the symmetry of the real part of $\phi_X(u)$ about zero the second term on the right hand side of Eqn (B.25) is zero. Noting that the the imaginary part of an even derivative of $\phi_X(t)$ is zero at $t = 0$, Eqn (B.25) becomes simply

$$(-1)^n \; \mathrm{E}\{|X|^{2n}\} = \frac{\partial^{2n} \phi_X(t)}{\partial t^{2n}} \mid_{t=0} \quad\quad (B.26)$$

as expected. Likewise

$$(-1)^n \; i\mathrm{E}\{|X|^{2n-1}\} = \frac{\partial^{2n-1} \mathrm{Re}\{\phi_X(t)\}}{\partial t^{2n-1}} - \frac{i}{\pi} \int_{-\infty}^{\infty} \frac{(2n-1)!}{(-u)^{2n}} \; \mathrm{Re}\phi_X(u) du \quad (B.27)$$

The first term on the right hand side of (B.27) is zero, therefore

$$\mathrm{E}\{|X|^{2n-1}\} = -(-1)^n \; \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{(2n-1)!}{u^{2n}} \; \mathrm{Re}\phi_X(u) du \quad\quad (B.28)$$

Since the imaginary part of $\phi_X(u)$ is odd we have as a final result

$$\mathrm{E}\{|X|^{2n-1}\} = -(-1)^n \; \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{(2n-1)!}{u^{2n}} \; \phi_X(u) du \quad\quad (B.29)$$

From Eq. (B.3), the expected value of the modulus of the G.C.C. trajectory, $|R^g(\tau)|$ is known if the characteristic function of $R^g(\tau)$ is known. Defining Z as

$$Z_{t,\tau} = |R^g(t) R^g(\tau)| \quad\quad (B.30)$$

Eq. (B.3) gives the autocorrelation function of $|R^g(\sigma)|$ if we know the two dimensional characteristic function of $R^g(t)$, $R^g(\tau)$. Define

$$\nu(t, \tau) = E\{Z_{t,\tau}\} - E\{\,|\,R^g(t)\,|\,\} E\{\,|\,R^g(\tau)\,|\,\} \qquad (B.31)$$
$$\Psi_t(u) = E\{e^{iuR'(t)}\}$$
$$\Phi_{t,\sigma}(u) = E\{e^{iuR'(t)R'(\tau)}\}$$

Then from (B.29),

$$\mathbf{E}\{\,|\,R^g(t)\,|\,\} = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{1}{u^2}\,\Psi_t(u)du \qquad (B.32)$$

$$\nu(t, \tau) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{1}{u^2}\,\Phi_{t,\tau}(u)du$$

$$- \left[\frac{1}{\pi} \int_{-\infty}^{\infty} \frac{1}{u^2}\,\Psi_t(u)du\right]\left[\frac{1}{\pi} \int_{-\infty}^{\infty} \frac{1}{u^2}\,\Psi_\tau(u)du\right]$$

Substituting the above equations into the expression Eq. (13) gives the result, Eq. (21).

Appendix C

Here the covariances of the sample cross-correlation $\hat{R}_{12}(\tau)$ and the GCC trajectory $\hat{R}^{g}_{12}(\tau)$, Eq. (2), are derived. It is then shown that the above two correlation functions are approximately covariance stationary outside of the region $[-T_c + D_0, T_c + D_0]$, where D_0 is the true delay parameter. We will specifically assume that the waveforms x_1 and x_2 are Gaussian random processes as given in Eq. (1) with spectra G_{11}, G_{22} and cross-spectrum G_{12}. Define T_c the signal correlation time.

Let $W(w) \equiv 1$ in the expression for the GCC trajectory, Eq. (2). Then the cross-correlation, $\hat{R}_{12}(\tau)$, is

$$\hat{R}_{12}(\tau) = \frac{1}{m} \sum_{k=1}^{m} \int_{T(k-1)}^{Tk} \frac{m}{T} \, x_{1k}(\tau) \, x_{2k}(\sigma+\tau) d\sigma \quad , \quad \tau \in [-D_m, D_m] \quad \text{(C.1)}$$

Define the covariance

$$\sigma_{\eta, \eta+\tau} = cov\left(\hat{R}_{12}(\eta), \hat{R}_{12}(\eta+\tau)\right) \quad \text{(C.2)}$$

Since $T/k \gg T_c$ the integrals in Eq. (C.1) are approximately independent and the covariance operator in Eq. (C.2) commutes with the summation in Eq. (C.1) to yield

$$\sigma_{\eta,\eta+\tau} = \frac{1}{T^2} \sum_{k=1}^{m} cov\left(\hat{R}^{k}_{12}(\eta), \hat{R}^{k}_{12}(\eta+\tau)\right) \quad , \quad \tau \in [-D_m, D_m] \quad \text{(C.3)}$$

where $\hat{R}^{k}_{12}(\eta) = \int_{T(k-1)}^{Tk} x_{1k}(\sigma) \, x_{2k}(\sigma+\eta) d\sigma$.

For simplicity, we will take m to be unity in Eq. (C.3), the result for the general case follows directly. Thus Eq. (C.3) becomes

$$\sigma_{\eta, \eta+\tau} = \int_{0}^{T} d\sigma \int_{0}^{T} du \, cov\left(x_{1k}(\sigma) \, x_{2k}(\sigma+\eta), \, x_{1k}(u) \, x_{2k}(u+\eta+\tau)\right) \quad \text{(C.4)}$$

From the Gaussian and zero mean assumptions imposed on x_1 and x_2, we can express the covariance under the sum in Eq. (C.4) as

$$cov\Big(x_{1k}(\sigma)x_{2k}(\sigma+\eta),\ x_{1k}(u)x_{2k}(u+\eta+\tau)\Big) = \qquad (C.5)$$

$$= E\{x_{1k}(\sigma)\ x_{1k}(u)\}\ E\{x_{2k}(\sigma+\eta)\ x_{2k}(u+\eta+\tau)\}$$
$$+ E\{x_{1k}(\sigma)\ x_{2k}(u+\eta+\tau)\}\ E\{x_{2k}(\sigma+\eta)\ x_{1k}(u)\}$$
$$= R_{11}(u-\sigma)\ R_{22}(\tau+u-\sigma) + R_{12}(\eta+\tau+u-\sigma)\ R_{12}(\eta-(u-\sigma)) \qquad (C.6)$$

by stationarity. Substitute (C.6) in (C.4) and make a change of variable $\xi = u-\sigma$ so that

$$\sigma_{\eta,\,\eta+\tau} = \frac{1}{T}\int_{-T}^{T}\Big[R_{11}(\xi)\ R_{22}(\xi+\tau) + R_{12}(\eta+\tau+\xi)\ R_{12}(\eta-\xi)\Big] \bullet \Big[1-\frac{|\xi|}{T}\Big]d\xi \qquad (C.7)$$

Now for a fixed η, take the Fourier transform of both sides of Eq. (C.7) to obtain after some manipulation

$$\int_{-\infty}^{\infty} \sigma_{\eta,\,\eta+\tau}\ e^{jw\tau}d\tau = \qquad (C.8)$$

$$= \frac{1}{T}\Big[G_{11}(w)\ G_{22}(w) + G_{12}^{2}(w)\ e^{j2\eta w}\Big] * sinc^2(wT/2)$$

If $T \gg T_c + D_m$, then the factor $[\,1-\frac{|\xi|}{T}\,]$ in (C.7) is approximately unity throughout the non-zero region of the integrand. In this case,

$$\sigma_{\eta,\,\eta+\tau} = \frac{1}{T2\pi}\int_{-\infty}^{\infty}\Big[G_{11}(w)\ G_{22}(w) + G_{12}^{2}(w)\ e^{j2\eta w}\Big]e^{jw\tau}dw \qquad (C.9)$$

For the general case, m greater than unity, if $T/m \gg T_c + D_m$ then from the stationarity of x_1 and x_2 the sum in Eq. (C.3) is over identical quantities given by Eq. (C.9). Therefore

$$\sigma_{\eta,\,\eta+\tau} = \frac{m}{T2\pi}\int_{-\infty}^{\infty}\Big[G_{11}(w)\ G_{22}(w) + G_{12}^{2}(w)\ e^{j2\eta w}\Big]e^{jw\tau}dw \qquad (C.10)$$

Let $s(t)$ be a bandlimited white noise of bandwidth $2\pi/T_c$. Then

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} G_{12}^2(w) \, e^{j2\eta w} \, e^{jw\tau} dw = K \, sinc(\pi(2(\eta - D_o) + \tau)/T_c) \qquad \text{(C.11)}$$

where K is a constant proportional to the signal power. Since $sinc(x)$ falls off rapidly for $|x| > T_c$ to a good approximation for η, $\eta + \tau$ outside of the interval $[D_o - T_c, D_o + T_c]$ we have

$$\sigma_{\eta, \eta + \tau} = \frac{m}{T2\pi} \int_{-\infty}^{\infty} G_{11}(w) \, G_{22}(w) \, e^{jw\tau} dw \qquad \text{(C.12)}$$

That is, R_{12} is stationary over the above region. Depending on one's definition of T_c, for a general signal, stationarity will hold for sufficiently large $|\eta|$.

It has been noted elsewhere [5] that for a general real GCC wieghting, $W(w)$, the GCC is equivalent to passing the waveforms x_1 and x_2 through filters H_1 and H_1' respectively, where $H_1 H_1' = W$, and cross-correlating the outputs. Therefore for the GCC implementation

$$\sigma_{\eta, \eta + \tau} = \frac{m}{T2\pi} \int_{-\infty}^{\infty} \left[G_{11}(w) \, G_{22}(w) + G_{12}^2(w) \, e^{j2\eta w} \right] |W(w)|^2 \, e^{jw\tau} dw \qquad \text{(C.13)}$$

where it is assumed that $T/m >> D_m + T_c'$ and T_c' is as defined in Section II, Eq. (3).

In the derivation of local variance, undertaken in Section IV, the signal $s(t)$ is assumed sufficiently broadband so that its trajectory can be taken as stationary except in the immediate neighborhood of the true delay parameter D_o. If $W(w)$ does not cause significant spreading of the signal auto-correlation function then we can approximate the covariance of R^g by

$$\sigma_{\eta, \eta + \tau} = \frac{m}{T2\pi} \int_{-\infty}^{\infty} G_{11}(w) \, G_{22}(w) \, |W(w)|^2 \, e^{jw\tau} dw \, , \qquad |\eta| > D_o \quad \text{(C.14)}$$

and for $\eta = D_o$

$$\sigma_{\eta, \eta + \tau} = \frac{m}{T2\pi} \int_{-\infty}^{\infty} \left[G_{11}(w) \, G_{22}(w) + |G_{12}(w)|^2 \right] |W(w)|^2 \, e^{jw\tau} dw \qquad \text{(C.15)}$$

Thus we can identify a GCC "spectrum"

$$C(w) = G_{11}(w)G_{22}(w) + |G_{12}(w)|1(\eta - D_o) \qquad (C.16)$$

here $1(\tau)$ is the indicator of $\tau = 0$.

REFERENCES

Chapter 3

1. C.H. Knapp and G.C. Carter, "The Generalized Correlation Method for Estimation of Time Delay," *IEEE Trans. on ASSP*, Vol. ASSP-24, No. 4, pp. 320-327, 1976.

2. S.K. Chow and P.M. Schultheiss, "Delay Estimation Using Narrowband Processes". *IEEE Trans. on ASSP*, Vol. ASSP-29, No. 3, pp. 478-484, 1981.

3. V.H. MacDonald and P.M. Schultheiss, "Optimum Passive Bearing Estimation in a Spatially Incoherent Noise Environment", *Journal of the Acoustical Society of America*, Vol. 46, No. 1, Part 1, pp. 37-43, 1969.

4. A.J. Weiss and E. Weinstein, "Fundamental Limitations in Passive Time Delay Estimation - Part I: Narrowband Systems," *IEEE Trans. on ASSP*, Vol. ASSP-31, No. 2, pp. 472-485, April 1983.

5. G.C. Carter, "Time Delay Estimation," Naval Underwater Systems Center, New London Lab, New London, CT, Tech Report 5335, 9 April 1976.

6. E.J. Hannan and P.J. Thomson, "Estimating Group Delay," *Biometrika*, vol. 60, No. 2, pp. 241-253, 1973.

7. A.O Hero and S.C. Schwartz, "The Center of Symmetry Estimator for Time delay," Submitted for publication in IEEE Trans on ASSP.

Very Narrowband GCC output.

Correct Peak Occurs at Do
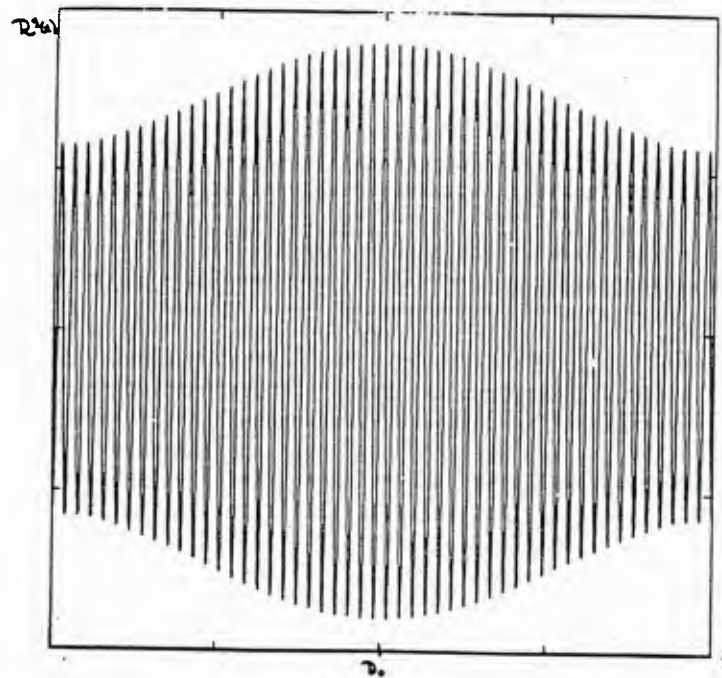
Fig. 2.1

A Small Noise Increase Causes Abrupt Jump in Peak Estimate
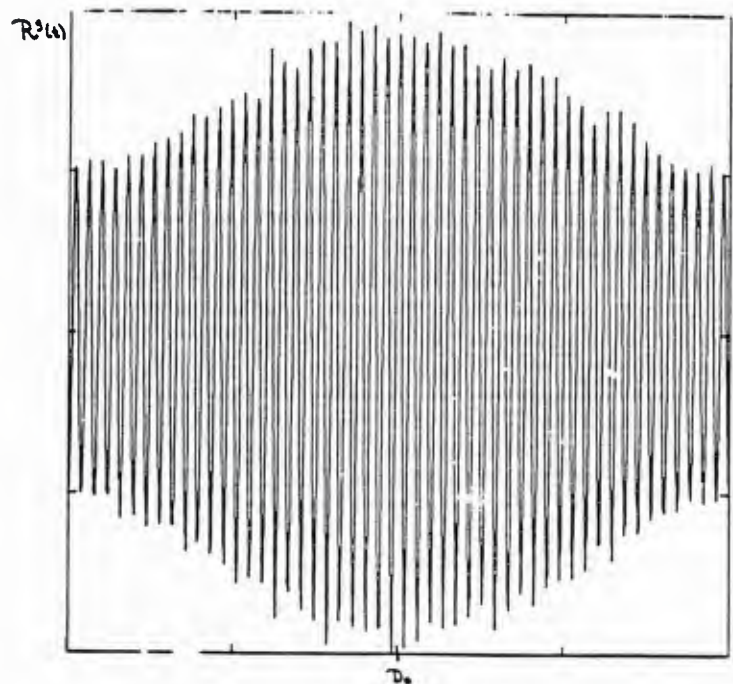
Fig. 2.2
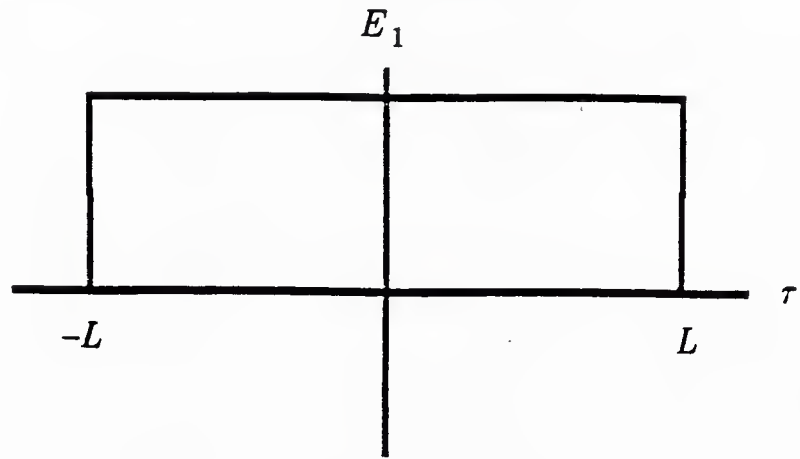
Fig. 3.1a
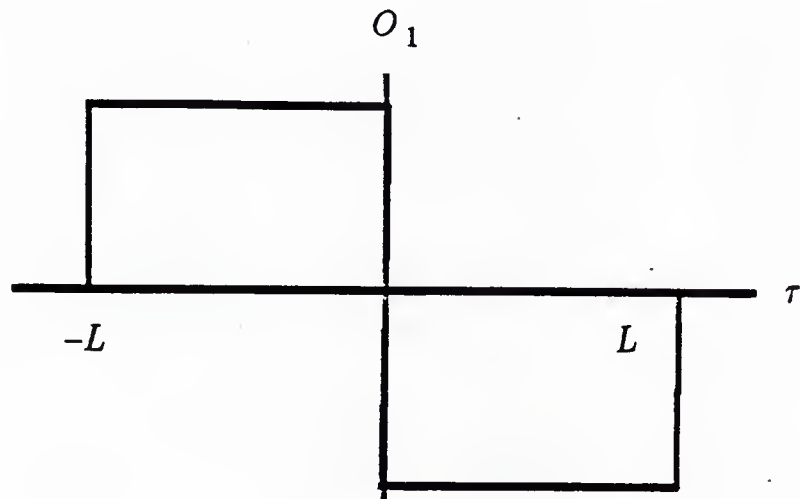
Fig. 3.1b
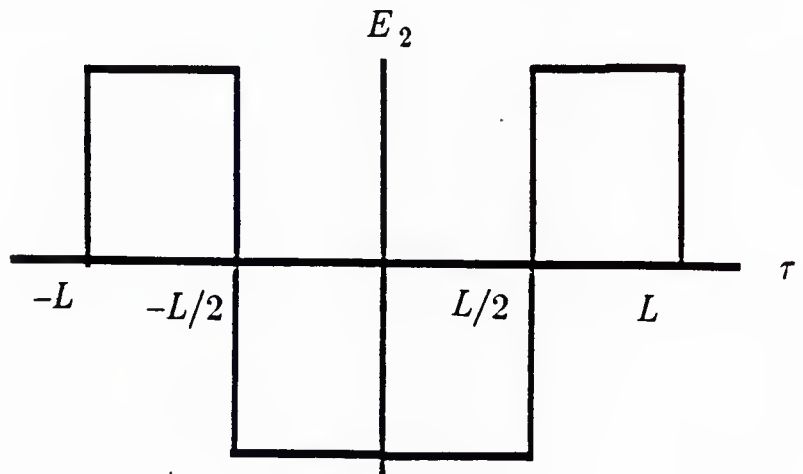
Fig. 3.1c

LCSE procedure compares area between '+' and '-'
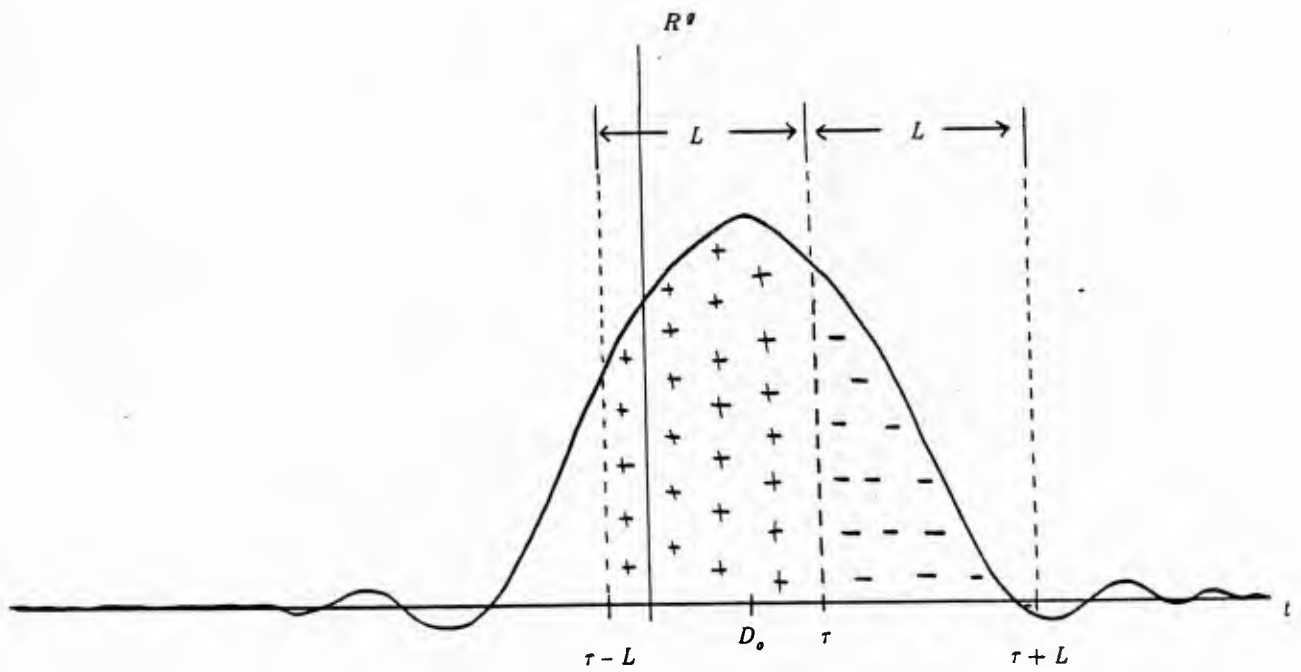
Fig. 3.2

Local error approximation
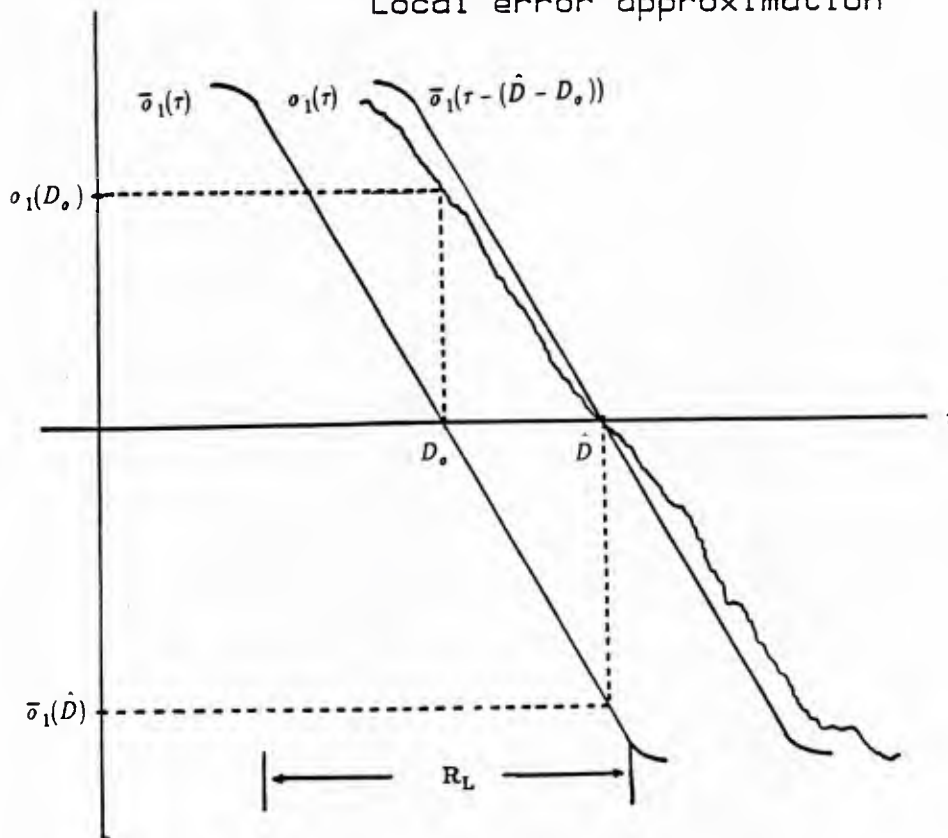
Fig. 4.1

Signal Spectrum

Frequency
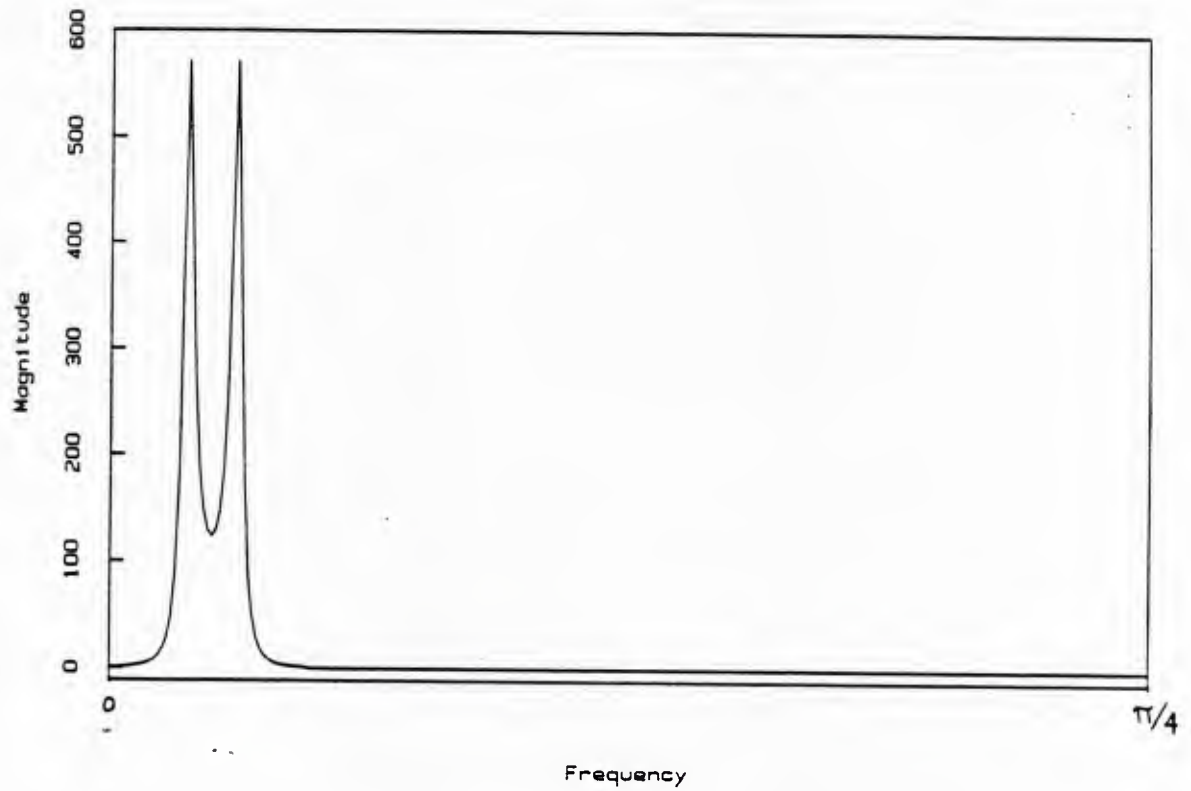
Fig. 5.1

Auto-correlation of signal z(k)

SNR (dB)
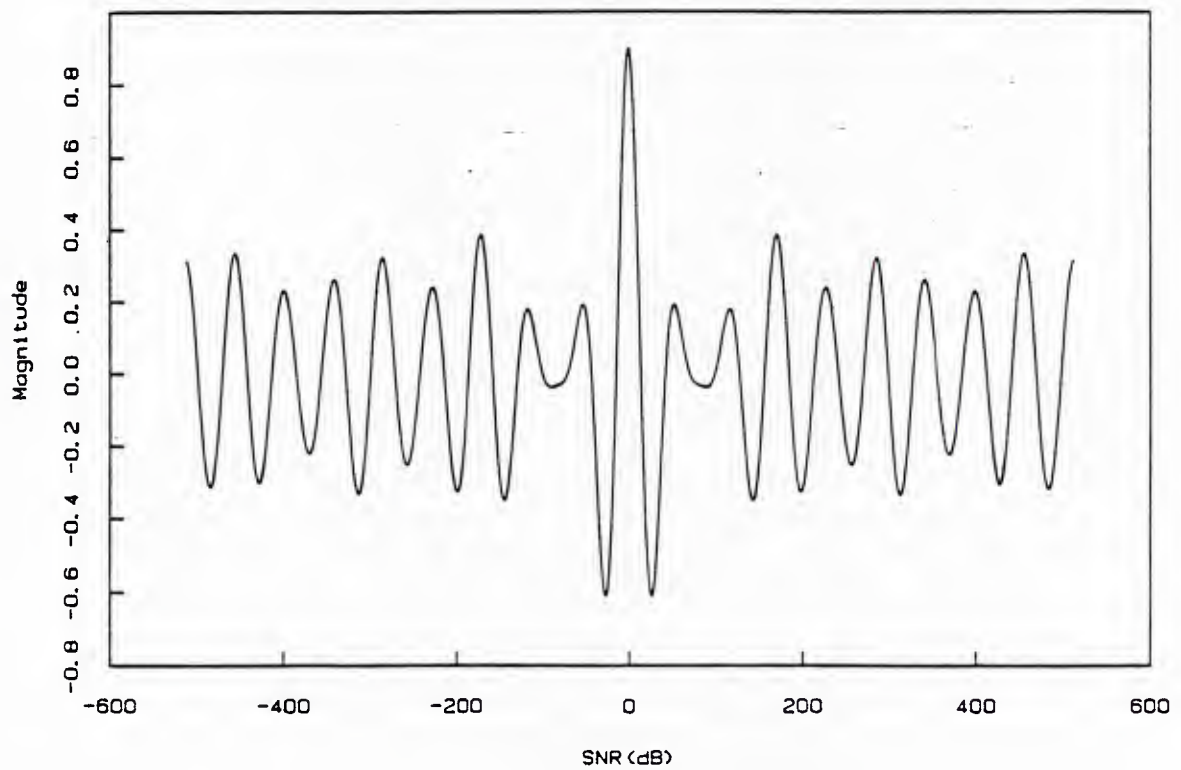
Fig. 5.2

Mean-Square-Error for Various Processors
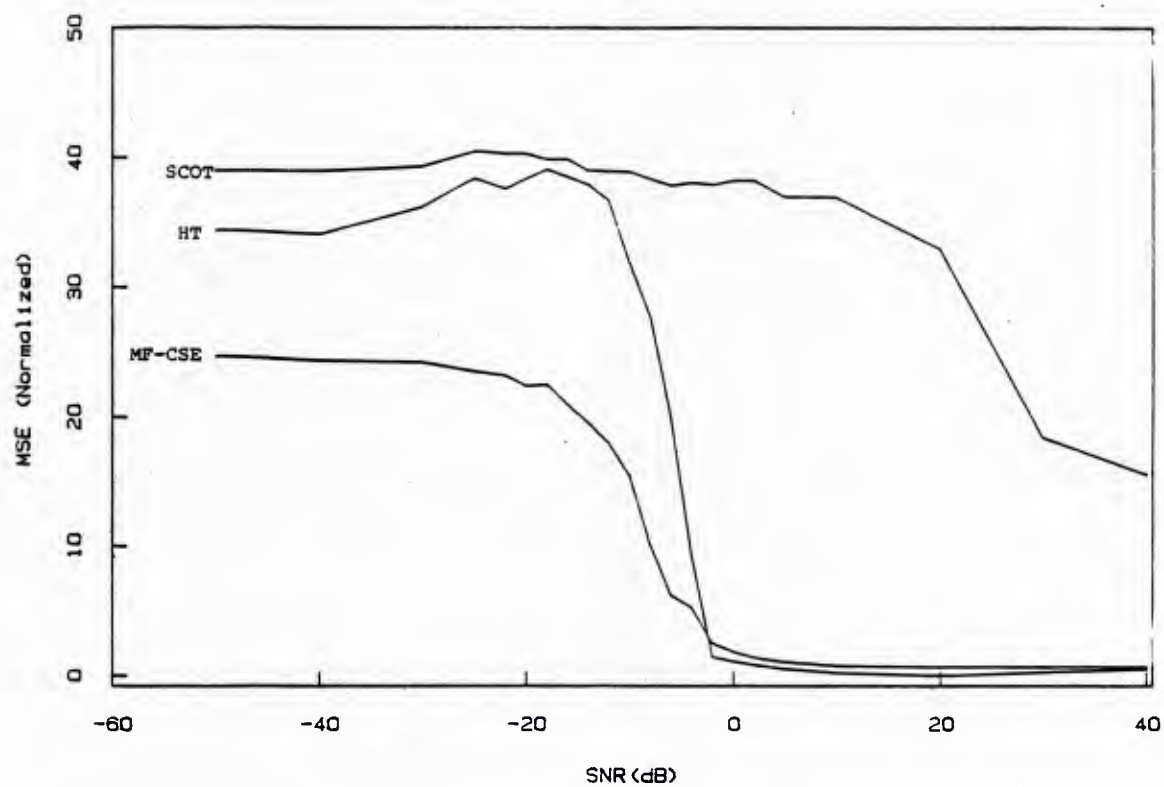
Fig. 5.3

Mean for Various Processors
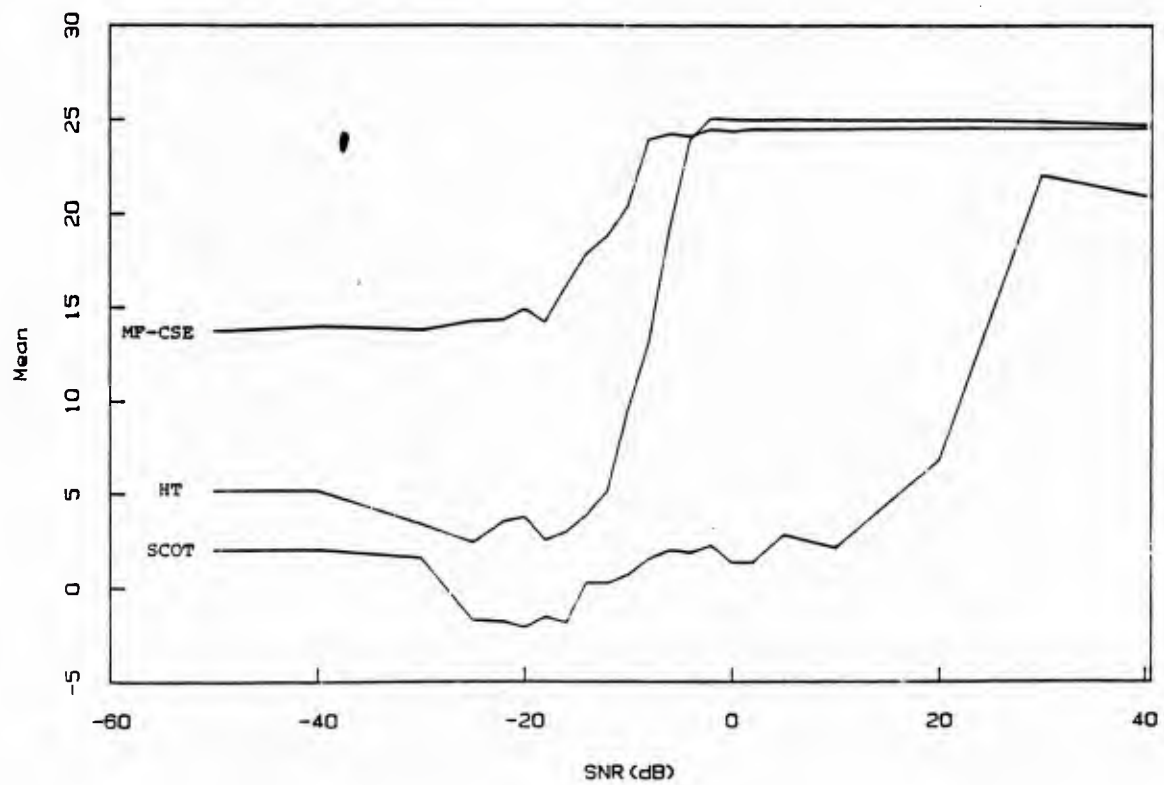
Fig. 5.4

Proportion of Estimates with 8% Accuracy
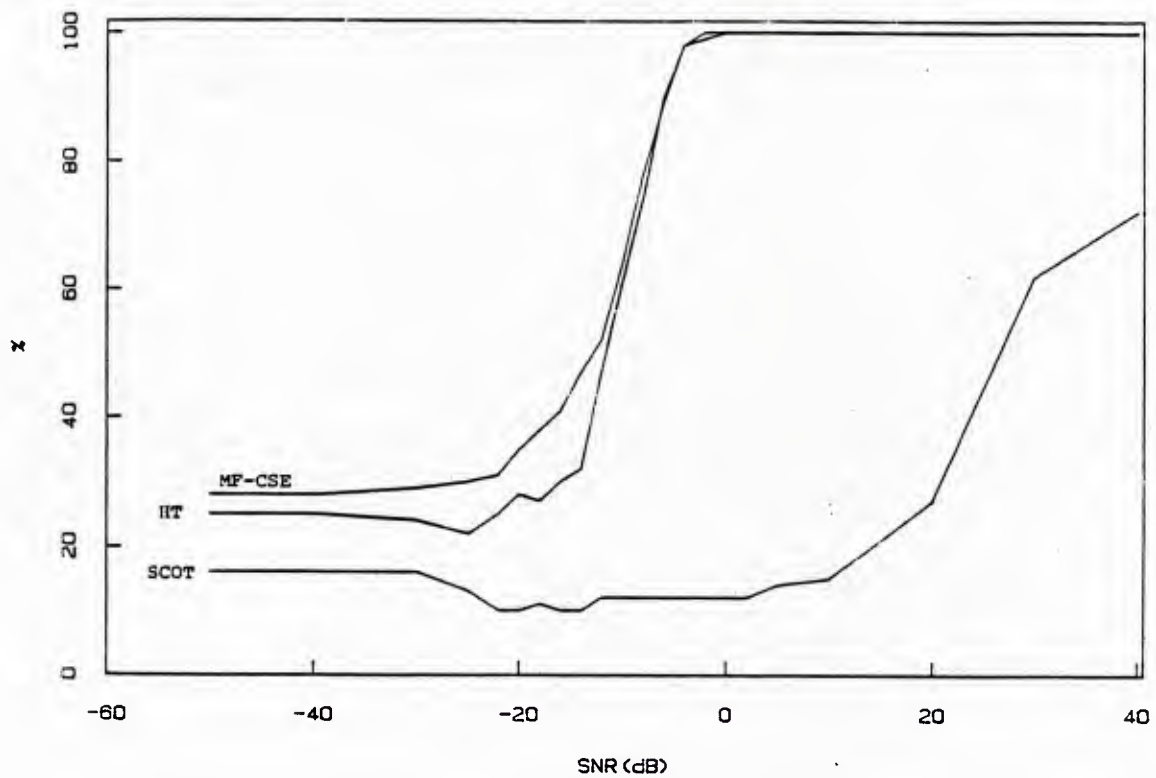
Fig. 5.5

Mean-Square-Error for Various Processors
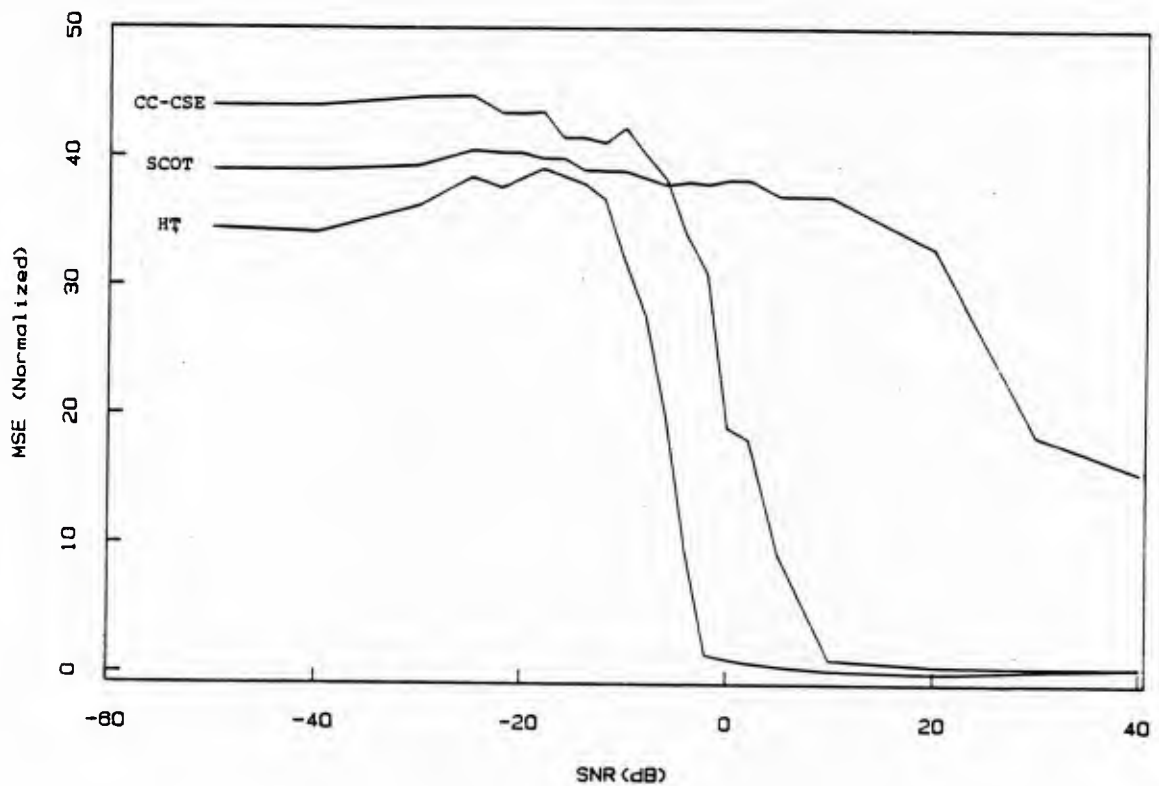
Fig. 5.6

Mean for Various Processors
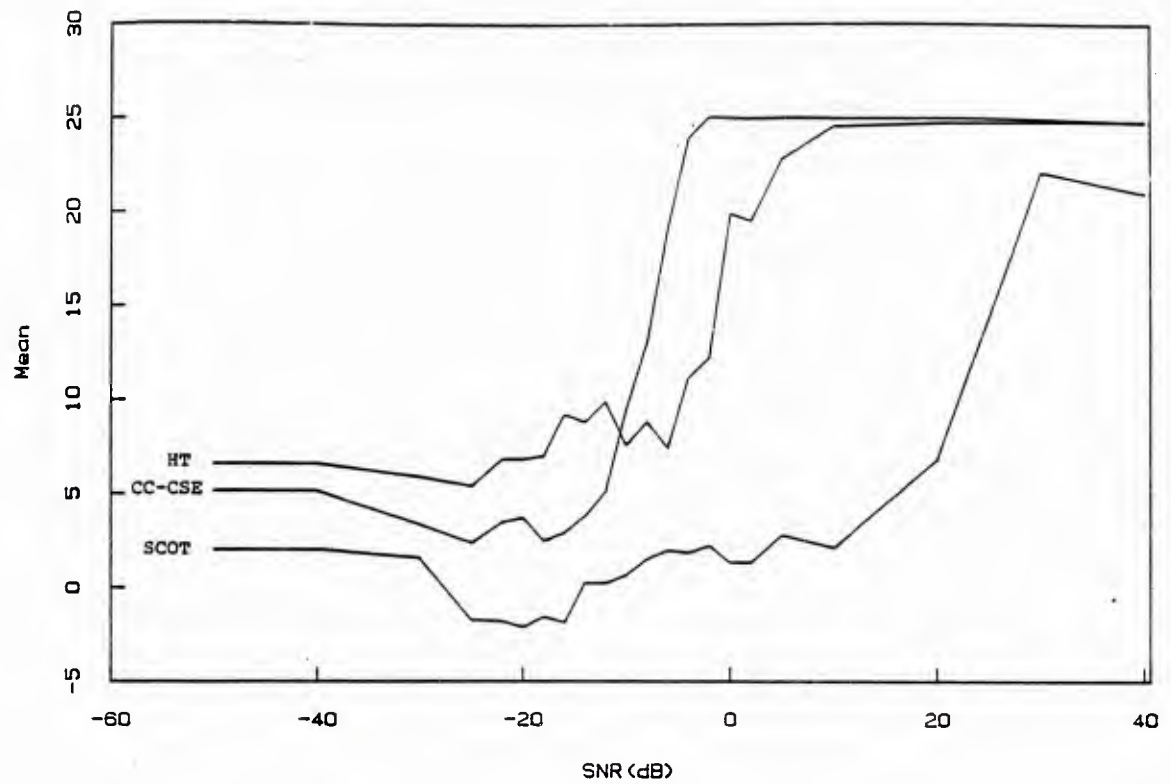
Fig. 5.7

Proportion of Estimates with 8% Accuracy
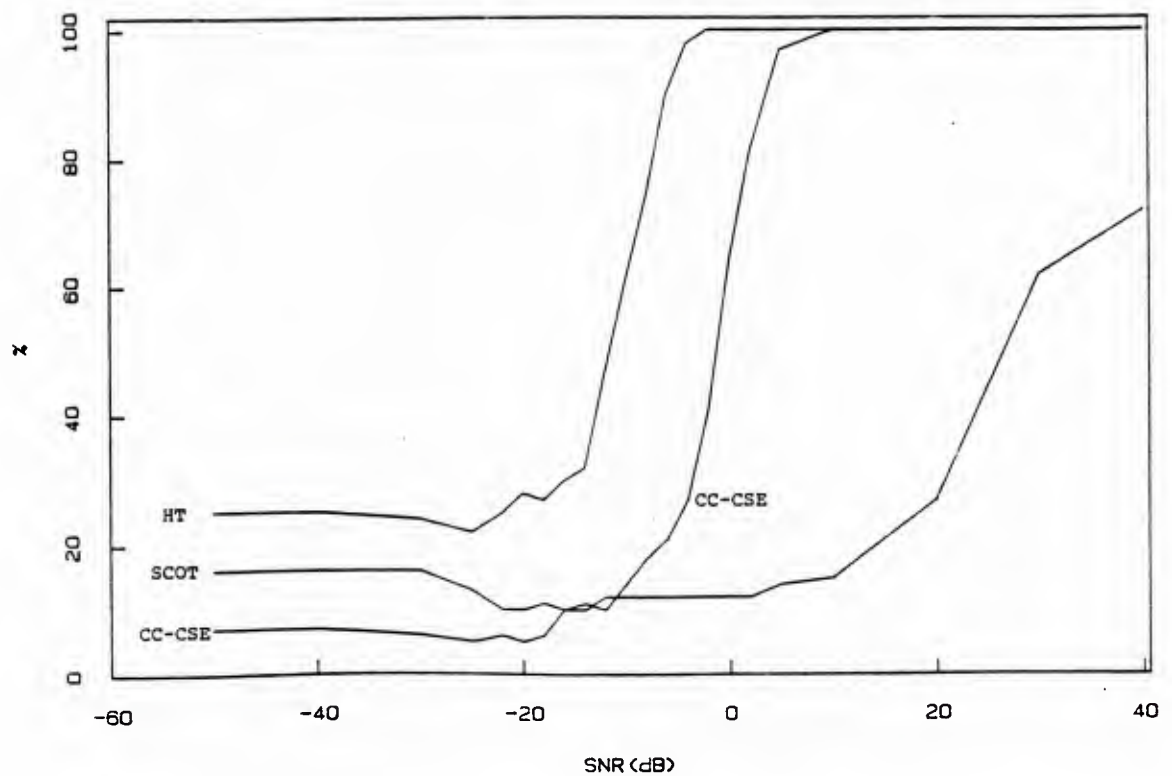
Fig. 5.8

Integration Regions

Fig. A.1

Fig. A.2

69
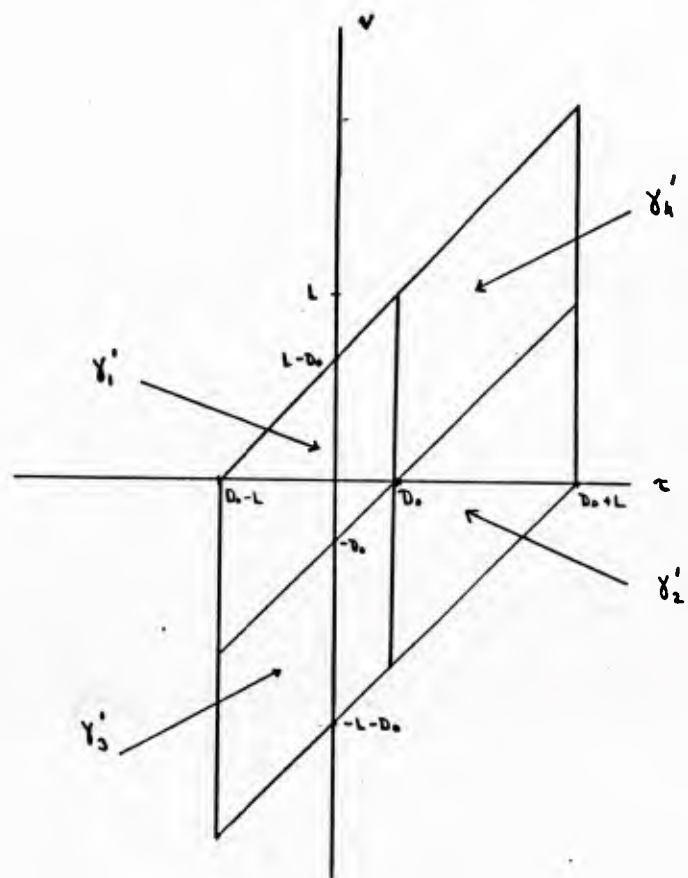
Chapter 4
On The Asymptotic Form of Level
Exceedance Probabilities

I. Introduction

The distribution function of the number of crossings of a level by a random process in a given interval is an important, but difficult to obtain, function which has received considerable attention in recent years [8]. Explicit results are known only for a handfull of specific random processes, e.g. the Markov class. In this chapter a general representation for the probability of getting one or more level upcrossings is obtained for a wide class of random processes, which allows the deviation of this probability from the probability of getting one or more points from an inhomogeneous Poisson process to be characterized. This representation is then exploited to show a limiting result analogous to results obtained by Leadbetter using extreme value theory for stationary processes [7]. Specifically, we assume certain asymptotic conditions on the trajectories of the process such as mixing. It is demonstrated that if the upcrossings of zero are made to become progressively rarer events, in a sense to be made clear later, then a normalized version of the number of upcrossings as a function of time converges in distribution to a Poisson process. Although the assumptions necessary for the proof reduce to somewhat stronger conditions on the random process than those of Leadbetter [7] for the stationary case, the result extends the asymptotic theory of level crossings to a certain subclass of non-stationary processes.

Results of the above type are of interest in connection with maximum likelihood parameter estimation when large errors may be significant. As is well known, the Cramer Rao Lower Bound only characterizes local error; that is, when the estimate is in the immediate vicinity of the true parameter [11]. When the trajectory of the so called "likelihood function" is predisposed to display multiple

widely separated maxima over the parameter space an additional large error measure is needed. One possible choice is the probability that a section of the trajectory exceeds a threshold, specified by the height of the trajectory at the true parameter. This probability can then be expressed within the framework of level crossing probabilities. Based on the results of this chapter, a Poisson approximation will be used in the next chapter to model the large error behavior of the maximum likelihood estimator for time delay.

II: A Representation for the Probability of Up-crossings.

Let (Λ, Φ, P) be a complete probability space and define the nested sequence of σ-fields : Φ_t , $t \in R$, ($\Phi_t \subset \Phi$, $\Phi_s \subset \Phi_u$ for $s < u$). We assume the Φ_t-measurable random process $X(t)$, $-\infty < t < +\infty$, to have the following properties: separability, almost sure (a.s.) sample function continuity, existence of the bivariate densities of $X(t)$ and $X(\tau)$, $f_{t,\tau}(y,z)$, for $t \neq \tau$. Additional properties will be imposed shortly.

We define an upcrossing analogously to Leadbetter in [2]. A realization of $X(t)$, $x(t)$, upcrosses zero in the interval $[\sigma,\nu)$ if there exists an open interval centered at some point $t_u \in (\sigma,\nu)$, $(t_u - \delta, t_u + \delta)$ say, over which $X(t) < 0$ to the left of t_u and $X(t) > 0$ to the right of t_u. We denote this occurence by the notation $A_{\sigma,\nu}$. Symbolically

$$(1)$$
$$A_{\sigma,\nu} \equiv \{w \in \Lambda| E\ t_u \in (\sigma,\ \nu),\ E\ \delta > 0\ ;\ s.t.\ X_{t_u - h} < 0 < X_{t_u + h},\ \text{for}\ 0 < h < \delta\}$$

This definition essentially excludes any "non-regular" upcrossings such as tangencies or non-smooth intersections of zero. It is shown in [2] that $A_{\sigma,\nu}$ is Φ_ν-measurable and that non-regular upcrossings are of zero probability. Hence definition (1) is sufficiently general for our purposes.

We further define the number of upcrossings by $X(\tau)$ in $[\sigma, \nu)$, denoted $N(\sigma, \nu)$, as the number of distinct points t_u at which upcrossings occur. More specifically we define: $N(\sigma,\nu) = \lim_{\Delta \to \infty} \sum_{t_i} I(A_{t_i, t_i + \Delta})$. Here $I(A)$ is the indicator of A and $\{t_i\}$ is an increasingly dense partition of (σ, ν) with inter grid spacing Δ.

By the assumed continuity of the random process $X(t)$ it is reasonable to expect that it can be well approximated by a piecewise linear process tied to $X(t)$ at a sufficiently dense set of points $t = t_o, t_1, ..., t_M$. That is let $\xi_n(t)$ denote a random process defined on an interval $[t_o, t_f]$ for which

$$\xi_n(t) = \begin{bmatrix} X(t) & t = t_k \\ X(t_k) + \dfrac{[X(t_{k+1}) - X(t_k)]}{(t_f - t_o)2^{-n}} (t - t_k) & t_k < t < t_{k+1} \end{bmatrix} \tag{2}$$

$$t_k = t_o + k2^{-n}(t_f - t_o), \quad k = 0, 1, 2, ..., 2^n$$

If N_n is the number of upcrossings of zero by $\xi_n(t)$ then the following is due to Ylvisaker [1].

Lemma 1.1

> *Let $N_n(t)$ be the number of upcrossings of zero by ξ_n in the interval $[t_o, t)$. Then $N_n(t)$ is monotonically non-decreasing in n and converges to $N(t)$, the number of upcrossings of zero by X in the same interval, with probability one as $n \to \infty$.*

From the above lemma it follows by monotone convergence that

$$P(N(t) \leq k) = \lim_{n \to \infty} P(N_n(t) \leq k), \quad \text{for } k = 0,1,2,... \tag{3}$$

Hence as far as the computation of upcrossing probabilities is concerned $\xi_n(t)$ and $X(t)$ can be used interchangeably in the sense of (3).

The following will be important for the upcoming development and are essentially Theorems (2.1) and (2.2) of [10].

Lemma 1.2

Let $[t_o, t_f]$ have the partition $\{t_i\}_{i=0}^{2^n}$. Then with $N(t)$ the number of upcrossings of zero by X in $[t_o, t)$ and $N_n(\tau, \sigma)$ the number of upcrossings by ξ_n in $[\tau, \sigma)$

$$\mathbf{E}[N(t_f)] = \lim_{n \to \infty} \sum_{i=0}^{2^n-1} P(A_{t_i,t_{i+1}}) = \lim_{n \to \infty} \sum_{i=0}^{2^n-1} P(N_n(t_i,t_{i+1}) > 0) \quad (4)$$

Define $g_{t,\tau}(y,z)$ the joint density of $X(t)$ and $[X(t+\tau) - X(t)]/\tau$. Then by elementary transformations

$$g_{t,\tau}(y,z) = \tau f_{t,t+\tau}(y,y+\tau z) \quad (5)$$

The following are essential to the development and are known as Leadbetter's conditions [2, Thm. 2]

$$g_{t,\tau}(y,z) \text{ is continuous in } t,y \text{ for each } \tau,z \quad (6)$$
$$g_{t,\tau}(y,z) \to p_t(y,z) \text{ as } \tau \to 0 \text{ uniformly in } t,y \quad (7)$$
$$g_{t,\tau}(y,z) \le l(z) \text{ for all } t,\tau,y,z \quad (8)$$

where

$$\int_0^\infty z l(z) dz < \infty$$

If the above three conditions hold then the following representation theorem holds for the probability of getting an upcrossing in $[t_o, t)$, $P(A_{t_o,t})$, here denoted $p(t)$.

Theorem 1.1

Suppose $X(\tau)$ has continuous sample functions with probability one and let the conditions (6) through (8) hold. Then the expected value of the number of upcrossings of zero by $X(\tau)$ in any finite interval $[t_o, t_f)$ is

finite and given by

$$\mathbf{E}[N(t_o,t_f)] = \int_{t_o}^{t_f} h(\tau)d\tau \tag{9}$$

$$h(\tau) = \int_0^\infty z \, p_\tau(0,z)dz$$

Furthermore the probability of getting at least one upcrossing of zero by $X(\tau)$ in $[t_o,t)$, $p(t)$, satisfies the relation

$$p(t) = \int_{t_o}^t h(\tau)(1-p(\tau))d\tau + Q(t) \tag{10}$$

where

$$Q(t) = \lim_{n\to\infty} \sum_{i=0}^{2^n-1} q(t_i) \tag{11}$$

$$\tag{12}$$

$$q(t_i) = P(N_n(t_i,t_{i+1})>0, N_n(t_i)=0) - P(N_n(t_i,t_{i+1})>0)P(N_n(t_i)=0)$$

Here $\{t_i\}_{i=0}^{2^n}$ is a partition of $[t_o,t]$

Eqn. (9) is obtained directly by modifying the proof of Theorem 2 of Leadbetter for downcrossings [2] to the case of upcrossings. The proof of the rest of Theorem 1.1 depends on a particular decomposition of the event that an upcrossing of zero by ξ_n occurs on $[t_o,t)$, which we denote $B_{t_o,t}$. If $N_n(t)$ is finite we can define $B_{\sigma,\nu}^1$: the event that the first instance of an upcrossing occurs in the subinterval $[\sigma,\nu)$ of $[t_o,t)$. That is

$$B_{\sigma,\nu}^1 \equiv B_{\sigma,\nu} \cap \bar{B}_{t_o,\sigma} \tag{13}$$

where we read this as: ξ_n first upcrosses in $[\sigma,\nu)$ if there is an upcrossing in $[\sigma,\nu)$ but none in $[t_o,\sigma)$.

We note the following two rather obvious properties of $B_{\sigma,\nu}^1$.

For $[\sigma,\nu)$ and $[s,t)$ disjoint

$$B_{\sigma,\nu}^1 \text{ and } B_{s,t}^1 \text{ are disjoint} \tag{14}$$

and

$$B_{t_o,t}^1 \equiv B_{t_o,t} \quad , t \in [t_o,t) \tag{15}$$

Eqns. (14) and (15) follow directly from the definition (13). The following proposition is central to the decomposition alluded to above.

Proposition 1.1

Let $A_{\sigma,\nu}$ denote an upcrossing of zero, and $A_{\sigma,\nu}^1$ the first instance of an upcrossing, by a random process $X(t)$ in $[\sigma,\nu)$; where $X(t)$ has absolutely continuous distributions. Then if the expected number of upcrossings of zero in $[t_o,t)$, $N(t_o,t)$, is finite the following equivalence holds with probability one

$$A_{t_o,t} \equiv \bigcup_{i=0}^{2^n-1} A_{t_i,t_{i+1}}^1 \tag{16}$$

where $\{t_i\}_{i=0}^{2^n}$ is a partition of $[t_o,t]$

Proof

Note that the number of upcrossings in $[t_o,t)$ is finite with probability one since

$$P(N(t_o,t) > k) \leq \sum_{i=k+1}^{\infty} P(N(t_o,t) = i) \tag{17}$$

$$\leq \sum_{i=k+1}^{\infty} iP(N(t_o,t) = i)$$

which must converge to zero as $k \to \infty$ by the finiteness of the mean number of upcrossings. Thus a "first instance of an upcrossing" is well defined. The inclusion "$\supset$" in (16) is trivial since any upcrossing in a subinterval of $[t_o,t)$ implies an upcrossing occurred in the entire interval. As for "$\subset$", if there is an upcrossing in $[t_o,t)$ it is either interior to one of the $[t_i,t_{i+1})$ or at one of the endpoints t_i, $i = 0, 1, \ldots, 2^n-1$ However from the absolute continuity of the distribution of $X(t)$, with respect to Lebesgue measure, this latter event has probability zero. Therefore the proposition follows.

Proof of Theorem 1.1

Partition $[t_o, t)$ into $2^n - 1$ subintervals of length $\Delta = (t_f - t_o)2^{-n}$ for $n = 0, 1, \ldots$. That is we have intervals $[t_i, t_{i+1})$ with $t_i = t_o + i\Delta$, $i = 0, 1, \ldots, 2^n$. Define $B_{\sigma,\nu}$, $\sigma, \nu \in \{t_i\}_{i=1}^{2^n}$, the event that the polygonal approximation, ξ_n upcrosses zero in $[\sigma, \nu)$, i.e. $N_n(\sigma, \nu) > 0$. Then from Proposition 1 and Eqn. (15) for $k = 2^n$ $(t_k = t)$

$$P(B_{t_o, t_k}) = P(\bigcup_{i=0}^{2^n-1} B_{t_i, t_{i+1}} \cap \overline{B}_{t_o, t_i}) = \sum_{i=0}^{2^n-1} P(B_{t_i, t_{i+1}} \cap \overline{B}_{t_o, t_i}) \tag{18}$$

Here we have used the disjointness property (14). Now add and subtract the product $P(B_{t_i, t_{i+1}})P(\overline{B}_{t_o, t_i})$ from each term under the sum (18)

$$P(B_{t_o, t}) = \sum_{i=0}^{2^n-1} [P(B_{t_i, t_{i+1}})P(\overline{B}_{t_o, t_i}) + q(t_i)] \tag{19}$$

where

$$q(t_i) = P(B_{t_i, t_{i+1}} \cap \overline{B}_{t_o, t_i}) - P(B_{t_i, t_{i+1}})P(\overline{B}_{t_o, t_i}) \tag{20}$$

as in the statement of the Theorem, Eqn. (12).

Now $B_{t_i, t_{i+1}}$ is equivalent to the event

$$B_{t_i, t_{i+1}} \equiv \{\xi_n(t_i) < 0 < \xi_n(t_{i+1})\} \tag{21}$$

Define

$$\eta_n(t_i) = \frac{\xi_n(t_{i+1}) - \xi_n(t_i)}{\Delta} \tag{22}$$

Combining Eqns. (21) and (22)

$$B_{t_i, t_{i+1}} = \{\xi_n(t_i) < 0 < \xi_n(t_i) + \Delta\eta_n(t_i)\} \tag{23}$$
$$= \{\xi_n(t_i) \in (-\Delta z, 0), \ \eta_n(t_i) = z > 0\}$$

Therefore by the definition of the joint density, $g_{t_i, \delta}$, of $\xi_n(t_i)$ and $\eta_n(t_i)$

$$P(B_{t_i, t_{i+1}}) = \int_0^\infty dz \int_{-\Delta z}^0 g_{t_i, \Delta}(x, z) dx \tag{24}$$

Now make a change of variable in the argument z of (24) and substitute the result back into Eqn. (19) to obtain

$$P(B_{t_o,t}) = \sum_{i=0}^{2^n-1} [\Delta \int_0^\infty dz \int_{-z}^0 g_{t_i,\Delta}(\Delta\, z,z) P(\bar{B}_{t_o,t_i}) dz + q(t_i)]$$

(25)

By the pointwise continuity and uniform convergence conditions, (6) and (7), for Δ sufficiently small

$$\int_{-z}^0 g_{t_i,\Delta}(\Delta z,z)dz = z\, \mathrm{p}_\tau(0,z) \ , \quad \tau \in [t_i,t_{i+1}]$$

(26)

Condition (8) asserts that $z\ell(z)$ is integrable over the positive real line where $\ell(z)$ upper bounds $g_{t,\tau}(y,z)$. Therefore the limit of (26) as $\Delta \to 0$ is bounded except possibly on some set of measure zero. From Lemma 1.1 and Eqn. (3) $P(\bar{B}_{t_o,t_i})$ converges to $P(\bar{A}_{t_o,t_i}) = 1 - p(t_i)$. Defining

$$a_\Delta(t_i,z) = \int_{-z}^0 g_{t_i,\Delta}(\Delta z,z) P(\bar{B}_{t_o,t_i}) dz$$

(27)

we have as Δ goes to zero

$$a_\Delta(t_i,z) \to z\, \mathrm{p}_{t_i}(0,z)(1 - p(t_i)) \qquad a.e.$$

(28)

and

$$a_\Delta(t_i,z) \le z\ell(z)$$

(29)

Hence by dominated convergence the first term of the expression (19) becomes in the limit

$$\lim_{n\to\infty} \sum_{i=0}^{2^n-1} P(B_{t_i,t_{i+1}}) P(\bar{B}_{t_o,t_i})$$

(30)

$$= \lim_{n\to\infty} \Delta \sum_{i=0}^{2^n-1} \int_0^\infty dz \int_{-z}^0 g_{t_i,\Delta}(\Delta\, z,z) P(\bar{B}_{t_o,t_i})$$

$$= \lim_{n\to\infty} \int_{t_o}^t d\tau \int_0^\infty dz \int_{-z}^0 dz \ g_{\tau,\Delta}(\Delta\, z,z) P(\bar{B}_{t_o,\tau})$$

$$= \int_{t_o}^t d\tau \int_0^\infty z\, \mathrm{p}_\tau(0,z)(1 - p(\tau)) dz < \infty$$

This is the first term in Eqn. (10).

From the expression (20)

$$-P(B_{t_i,t_{i+1}})P(\bar{B}_{t_o,t_i}) \leq q(t_i) \leq P(B_{t_i,t_{i+1}})P(B_{t_o,t_i}) \qquad (31)$$

so that the $q(t_i)$ are absolutely summable by Lemma 1.2 and the finiteness of $E[N(t_o,t_f)]$. Finally Theorem 1.1 follows by performing the limiting operation in (16) as $n \to \infty$, taking account of the regularity conditions shown above.

III. Asymptotic Results Concerning the Nature of Upcrossings

Theorem 1.1 gives an implicit relation for the probability of getting an upcrossing in a bounded interval. Although the rate function $h(\tau)$ may be known, in general the $Q(t)$ term in Eqn. (11) involves quantities which are not known. Essentially $Q(t)$, or $q(t_i)$, $i = 0, 1, \ldots, 2^n - 1$, (Eqns. (11) and (12)), are measures of the dependency structure of the upcrossing process N over disjoint intervals. Hence it is conceivable that Eqn. (11) may lend itself to an iterative approximate solution if some sequence of increasingly good estimates of the $Q(t)$ term is available. For the present, however, Theorem 1.1 can only be used as a verification method for some upcrossing probability candidate, $p(t)$, obtained perhaps by guessing. No constructive procedure for obtaining $p(t)$ is offered here.

Eqn. (11) can be used to prove certain asymptotic results for a fairly general class of upcrossing processes, which we will now undertake to show. To motivate these results the following argument is useful. Refering to Eqn. (11), assume that N_n converges to an independent increment point process $\tilde{N}$ as $n \to \infty$. Then $q_n(t_i)$ converges to zero for all i and by Eqn. (31), Lemma 1.2 and the finiteness of the mean number of upcrossings, $q_n(t_i)$ is summable over i as the t_i become dense in $[t_o,t)$. Dominated convergence then allows us to assert that $Q(t) = 0$ and Eqn. (11) becomes equivalent to a linear, first order, homogeneous

differential equation with coefficient $h(t)$ and initial condition $p(t_o) = 0$. This then has the solution

$$p(t) = 1 - exp(-\int_{t_o}^{t} h(\tau)d\tau) \qquad (32)$$

Eqn. (32) is of course valid for any semiclosed subinterval of $[t_o, t)$. Hence, by the independence of N over disjoint intervals, the upcrossing process must actually be an inhomogeneous Poisson process with rate parameter $h(\tau)$.

Unfortunately the above argument is fallacious since, roughly speaking, for non-zero N on bounded intervals the independent increment property of N is incompatible with the sample function continuity of $X(t)$ so that Theorem 1.1 does not even apply. Clearly then the pointwise convergence of N_n to an independent increment process $\tilde{N}$ is an overly strong imposition on $X(t)$. However in the following it will be shown that for a sequence of "thinned out" upcrossing count processes, $N^m(t_o, t)$ associated with $X(t)$, a related (time normalized) random process can be defined which converges weakly to a Poisson random process defined on the interval $[0,1)$ as $m \to \infty$. These results will depend on additional assumptions, such as mixing, on the distributions of $X(t)$.

The basic idea is to coordinate sequences of random processes $\{X_m\}$ and intervals $\{I_m\}$ such that upcrossings of zero by X_m become "rare events" while maintaining a non-zero expected value of the total number of upcrossings over I_m. For I_m sufficiently large any pair of upcrossings will become approximately independent since, with probability close to one, the events are separated in time by an amount exceeding the "inter-dependence time" (correlation time for Gaussian case) of X_m which can be specified by the mixing condition. Then Theorem 1.1 can be used to give the solution (32).

For simplicity, and without weakening the results, we set t_o in Eq. (32) to

zero. Identify I_m with the scalar T_m

$$I_m = [0, T_m), \quad I_o = [0,1) \tag{33}$$

where T_m is strictly increasing to infinity as $m \to \infty$. In general, when there is multiple indexing, subscripts indicate indexing with respect to the partition, $\{t_i\}$, of the time interval under consideration and superscripts index the quantity with respect to the infinite sequences $\{X_m\}$ and $\{I_m\}$. Thus $N_n^m(t)$ denotes the number of upcrossings of zero by the polygonal approximation to X_m, ξ_n^m, over the interval $[0,t) \in I_m$. Likewise N^m is the number of upcrossings associated with X_m itself. Analogously to the development of Theorem 1.1 define $\Phi_{\sigma,\nu}^m$, the σ-field generated by X_m on $[\sigma,\nu)$; B_{t_i,t_j}^m, the event $N_n^m(t_i,t_j) > 0$, where t_i and t_j are points contained in the 2^n-th order partition of I_m; and $p_m(t)$, the probability that $X_m(t)$ upcrosses zero on $[0,t) \subseteq [0, T_m)$.

The next section is concerned with the various conditions which will be imposed on X_m for asymptotic independence of the trajectories and Poisson-like behavior of the upcrossings.

Asymptotic Conditions

A mixing condition is a statement concerning the asymptotic independence of the trajectories of a random process on disjoint intervals $[\sigma,\nu)$ and $[s,\tau)$ as $|s - \nu| \to \infty$. For example $X(t)$ is "strong mixing" [3] if

$$\sup_\tau |P(A \cap B) - P(A)P(B)| \le \beta_l \tag{34}$$

where A and B are arbitrary events

$$A \in \Phi_{\tau,\infty}, \quad B \in \Phi_{-\infty,\tau-l}$$

and

$$\lim_{l \to \infty} \beta_l = 0$$

The major weakness of "strong mixing" is that (34) becomes vacuous if either A or B are of vanishingly small probability. Indeed in the present context the event A will be contained in the event that an upcrossing of zero occurs in an exceedingly small interval which of course has exceedingly small probability. The needed condition here is the summability to zero of the differences below

$$\lim_{m,n\to\infty} \sum_{i=0}^{2^n-1} |P(A_i^m \cap B_i^m) - P(A_i^m)P(B_i^m)| = 0 \qquad (35)$$

$$A_i^m \in \Phi_{-\infty,t_i-l_m}^m \ , \quad B_i^m \in \Phi_{t_i,t_{i+1}}^m$$

$$l_m \to \infty \ , \quad l_m = o(T_m) \ \text{ as } m \to \infty$$

$$\{t_i\}_{i=0}^{2^n} \ , \quad \text{an increasingly dense partition of } [0,T_m)$$

A sufficient condition for (35), if the quantities $P(B_i^m)$ are summable over i, is the following form of so called "uniform mixing" [3].

Mixing Condition

With $\Phi_{\sigma,\nu}^m$ the σ-field generated by the trajectories of X_m in $[\sigma,\nu)$, T_m a monotonic sequence increasing to infinity, X_m is said to be uniform-asymptotically mixing (u-a mixing) if

$$|P(B_i^m) - P(B_i^m|A_i^m)| < \alpha_{m,l_m} \qquad (36)$$

where

$$A_i^m \in \Phi_{-T_m,t_i-l_m}^m \ , \quad B_i^m \in \Phi_{t_i,t_{i+1}}^m$$

and

$$\lim_{m\to\infty} \alpha_{m,l_m} = 0$$

with

$$l_m \to \infty, \ l_m = o(T_m) \ , \quad \text{as } m \to \infty$$

$$\{t_i\}_{i=0}^{2^n} \ , \quad \text{an increasingly dense partition of } [0,T_m)$$

Note that for a dense partition $\{t_i\}$ the conditioning in (36) will be on the zero proba-

bility event $X(\tau) = 0$ at some specific point τ, viewed as a limit through a horizontal window [6]. Thus in the limit of dense partitions, although the conditional probability may not be well defined in the conventional sense, (36) is well defined. We state the following lemma which generalizes the uniform mixing condition to multiple events.

Lemma 2.1

Assume that $X_m(t)$ is uniform mixing in the sense of (36). Fix $l > 0$ and for $r > 1$ let $E_1, E_2, \ldots, E_r$ be disjoint intervals indexed in increasing order, that is, $\sup\{\tau \in E_{i-1}\} < \inf\{\tau \in E_i\}$ for $i = 1, 2, \ldots, r$, and seperated by at least l. Then for $A_i^m \subset \Phi_{E_i}^m$

$$|P(\bigcap_{i=1}^{r} A_i^m) - \prod_{i=1}^{r} P(A_i^m)| < \alpha_{m,l_m} \sum_{i=2}^{r} P(A_i^m) \qquad (37)$$

Proof

For $r = 2$ Eqn. (37) follows directly from the definition Eqn. (36) and the identity $P(A, B) = P(A \mid B)P(B)$. We proceed by induction.

$$|P(\bigcap_{i=1}^{r} A_i^m) - \prod_{i=1}^{r} P(A_i^m)|$$

$$\leq |P(\bigcap_{i=1}^{r-1} A_i^m \cap A_r^m) - P(\prod_{i=1}^{r-1} A_i^m)P(A_r^m)| + P(A_r^m)|P(\bigcap_{i=1}^{r-1} A_i^m) - \prod_{i=1}^{r-1} P(A_i^m)|$$

$$\leq \alpha_{m,l_m} P(A_r^m) + \alpha_{m,l_m} \sum_{i=2}^{r-1} P(A_i^m)$$

by Eqn. (36) and the induction hypothesis.

In order to make the upcrossings exceedingly rare events as $m \to \infty$ the following "rarefaction" condition is used

Rarefaction Condition

With $N_n^m(\sigma, \nu)$ the upcrossings of zero by the polygonal approximation ξ_n^m in $[\sigma, \nu) \subset [0, T_m)$ N_n^m satisfies a rarefaction condition if for $l_m \to \infty$,

$$l_m \,=\, o(T_m)$$

$$\lim_{m \to \infty} \sum_{t_i = l_m}^{T_m} P(N_n^m(t_i, t_{i+1}) > 0, \, N_n^m(t_i - l_m, t_i) > 0) = 0 \qquad (38)$$

$\{t_i\}_{i=0}^{2^n}$, $n = n(m)$, *an increasingly dense partition of* $[0, T_m)$

The above condition is a strong condition on the trajectories similar to , but possibly more restrictive than, the condition $D_c{}'$ used in [7] for the stationary case. An additional condition needed is the following which is analogous to condition (4.6) in [7]

$$\frac{P(N_n^m(t, t+h) > 0)}{\mathbf{E}[N_n^m(t, t+h)]} \,\to\, 1 \quad as \;\; m \to \infty \qquad (39)$$

$\{t_i\}_{i=0}^{2^n}$, $n = n(m)$, *an increasingly dense partition of* $[0, T_m)$

for some h_o , $0 < h < h_o$ *and for all* $t \in [0, T_m)$.

Condition (39) is stronger than a well known necessary condition for a process to be (asymptotically) Poisson: for infinitesimal intervals the probability of getting a point is proportional to the expected number of points in the interval (linear in the length of the interval for stationary processes). The condition can be interpreted as an extension of this necessary condition to certain finite intervals.

We state here two additional conditions particular to the nonstationary situation, which guarantee that the behavior of the upcrossing process is sufficiently uniform over time to exclude degeneration of the upcrossing probabilities.

Uniform Denseness Condition

Let $N_n^m(\tau, s)$ *be the number of upcrossings of zero by* ξ_n^m, *in the interval* $[\tau, s)$. *Assume one is given a positive integer* K, *and an interval* $J = [\tau, s] \subset [0, \, T_m)$, $\tau, s \in \{t_i\}_{i=1}^{2^n}$, *whose length* $\|J\|$ *is* $o(\dfrac{T_m}{K})$. *Then*

given any point t_p in the partition $\{t_l\}_{l=1}^{2^n}$, and an $\epsilon > 0$ the Uniform Denseness Condition is satisfied if there exist sufficiently large M, M' and an interval $I = [\sigma, \nu] \subset [0, T_m)$, $\sigma, \nu \in \{t_i\}_{i=1}^{2^n}$, containing t_p whose length $||I||$ is also $o(\dfrac{T_m}{K})$ such that

$$|E[N_n^m(J)] - E[N_n^m(I)]| < \epsilon, \qquad m > M, n > M' \qquad (40)$$

Asymptotic Uniform Negligibility

Let N_n^m be the number of upcrossings of zero by ξ_n^m in $I_m = [0, T_m)$ and let $\{\tau_i\}_{i=0}^{K}$ be a partition of I_m of size K. Then with $N_{n_k}^m$ the number of upcrossings within the k-th partition element, $[\tau_k, \tau_{k+1}]$, X_m satisfies uniform asymptotic negligibility if

$$\lim_{m, K \to \infty} \frac{E[N_{n_k}^m]}{\sum\limits_{k=1}^{K} E[N_{n_k}^m]} = 0 \qquad (41)$$

Loosely speaking (40) implies that the upcrossings are lean enough so that "similar" intervals, of similar order with respect to I_m, have associated with them a "similar" expected number of upcrossings. This will be seen to imply a continuity property on $P(N^m(\tau, \sigma) > 0)$ viewed as a map from the sets $[\tau, \sigma)$ to the positive reals. (41) guarantees that in no case will the total number of upcrossings over I_m be dominated by upcrossings in small subintervals of I_m. If the process $X(t)$ were stationary these two conditions, (40) and (41), would be trivially satisfied since the expected values of $N_n^m(J)$ and $N_n^m(I)$ are identical if $J = [\tau, s)$ and $I = [\sigma, \nu)$ are intervals of equal length.

Main Theorem

With the above conditions we are prepared to state the main result concerning the convergence of a certain normalized upcrossing process, associated with $X_m(t)$, to a Poisson process.

Theorem 2.1

Let the a.s. continuous processes $X_m(t)$ have absolutely continuous distributions for $m = 1, \ldots$. Assume each X_m satisfies Leadbetter's conditions (6) through (8), with, in addition, $l^m(z) = O(T_m^{-1})$ in (8), u-a mixing of the form (36) and conditions surrounding Eqns. (38) through (41). Denote by $h_m(\tau)$, $\tau \in [0, T_m)$, the rate function of N^m, the upcrossings of zero by X_m, and assume that $T_m \cdot h_m(\tau T_m)$ converges uniformly in $\tau \in [0,1)$ to a function $h(\tau)$ as $m \to \infty$. Then with the normalized counting process $\tilde{N}^m(\tau) = N^m(\tau T_m)$ we have

$$\tilde{N}^m(\tau) \to N^*(\tau) \quad \text{in distribution} \tag{42}$$

where $N^*(\tau)$ is a non-stationary Poisson random process on $[0,1)$.

We will need the following lemma in order to use Theorem 1.1 for the proof of Theorem 1.2.

Lemma 2.2

Let $X(\tau)$ be a random process which satisfies the assumptions of Theorem 1.1 and denote by $N(\tau)$ the number of upcrossings of zero by X in $[0,\tau)$. Let $N^*(\tau)$ be the number of points generated by a Poisson process in $[0,\tau)$ with rate function $h(\sigma)$. Then the following inequality holds on the difference between $p(t) = P(N(t) > 0)$ and $p^*(t) = P(N^*(t) > 0)$.

$$|p(t) - p^*(t)| \leq \lim_{n \to \infty} \sum_{i=0}^{2^n - 1} |q(t_i)| \, exp(-\int_0^t h(\tau)d\tau) \tag{43}$$

Where $q(t_i)$ is given by Eqn. (12) and $\{t_i\}_{i=0}^{2^n}$ is an increasingly dense partition of the interval $[0,t)$.

Proof

Consider the following integral equation for $0 < t' < t$

$$p^*(t') = \int_0^{t'} h(\tau)(1 - p^*(\tau))d\tau \tag{44}$$

The right hand side of Eqn. (44) is the first term in Theorem 1.1, Eqn. (10), for $t = t'$ and has the solution

$$p^*(t') = 1 - \exp(-\int_0^{t'} h(\tau)d\tau) \tag{45}$$

which corresponds to the case where N^* is Poisson. Therefore subtracting Eqn. (44) from Eqn. (10) in Theorem 1.1 and using the triangle inequality on the resultant difference we have

$$|p(t') - p^*(t')| \le \int_0^{t'} h(\tau)|p(\tau) - p^*(\tau)|d\tau + |Q(t')| \ , \ \ t' \in [0,t] \tag{46}$$

Recall from Eqn. (11) if $\{t_i\}_{i=0}^{2^n}$ is a partition of $[0,t]$ with inter-grid spacing Δ

$$Q(t) = \lim_{\Delta \to 0} \sum_{i=0}^{\frac{t}{\Delta}-1} q(i\Delta) \tag{47}$$

where Δ goes to zero through the values $t2^{-n}$, $n = 0, 1 \dots$. From the triangle inequality

$$|Q(t)| \le \lim_{\Delta \to 0} \sum_{i=0}^{\frac{t}{\Delta}-1} |q(i\Delta)| \tag{48}$$

Initially assume that t' is contained in the set of points partitioning $[0,t)$ for sufficiently large n. With this assumption the sum in Eqn. (48) upper bounds $|Q(t')|$ for n sufficiently large since this latter quantity is upper bounded by the sum of a subset of the positive terms to the right of inequality (48). Therefore

$$|p(t') - p^*(t')| \le \int_0^{t'} h(\tau)|p(\tau) - p^*(\tau)|d\tau + r(t) \ , \ \ t' \in [0,t) \tag{49}$$

where

$$r(t) = \lim_{n \to \infty} \sum_{i=0}^{2^n - 1} |q(t_i)| \tag{50}$$

We can now apply the *Bellman - Gronwall* inequality for integral inequalities [4] to (49) to obtain

$$| p(t') - p^*(t')| \le r(t) \exp(\int_0^{t'} h(\tau) d\tau) \ , \quad t' \in [0, t) \tag{51}$$

From Eqn. (32) it is clear that $p^*(t)$ is a continuous function. This is also true for $p(t)$ since by the definition of an upcrossing, Eqn. (1), given $\epsilon > 0$ there must exist a $\delta < 0$ such that

$$0 \le P(N(t+\delta) > 0) - P(N(t) > 0) = P(N(t,t+\delta) > 0, N(t) = 0) \tag{52}$$
$$\le P(|X(t)| < \epsilon)$$

This last term can be made arbitrarily small by the absolute continuity of the distributions of X. Therefore Eqn. (51) must hold at $t' = t$.

If t' is not one of the grid points $\{k2^{-n} t\}_{k=0}^{2^n}$ for some n then for a given n let t_{l_0} be the closest grid point to the left of t'. Then for the partitions $\{k2^{-n} t, t'\}_{k=0}^{l_0}$ Theorem 1.1 holds for $p(t')$ with

$$Q(t') = \lim_{n \to \infty} \left\{ \sum_{k=0}^{l_0} q(t_k) \right.$$
$$\left. + [P(A_{t_{l_0}, t'}, \overline{A}_{0, t'}) - P(A_{t_{l_0}, t'}) P(\overline{A}_{0, t'})] \right\} \tag{53}$$

with $q(t_i)$ as in Eqn. (12) and $t_i = i2^{-n} t$. Now the difference in brackets displayed in Eqn. (53) becomes arbitrarily small as n becomes large. Hence the assumption that t' is a grid point entails no loss in generality and the lemma follows.

We would like Theorem 1.1, and thus Lemma 2.2, to hold for the sequence X_m having the properties indicated in Theorem 2.1. The following lemma addresses this.

Lemma 2.3

If the a.s. continuous processes $X_m(t)$ satisfies Leadbetter's conditions with,

$$l(z) = l^m(z) = O(T_m^{-1}) \tag{54}$$

in (8), and

$$T_m h_m(\tau T_m) \rightarrow h(\tau) \ , \quad as \ m \rightarrow \infty \tag{55}$$

where convergence is uniform in $\tau \in [0,1)$, then Theorem 1.1 holds for $[t_o,t_f) = [0,T_m)$. Furthermore with $N^m(t)$ the number of zero upcrossings by X_m on $[t_o,t)$

$$\mathbf{E}[N^m(T_m)] \rightarrow \int_0^1 h(t)dt < \infty \tag{56}$$

Proof

From Lemma 1.2 and the application of the arguments leading to Eqn. (24) in the proof of Theorem 1.1 we have the relation

$$\mathbf{E}[N^m(T_m)] = \lim_{\Delta \rightarrow 0} \int_0^{T_m} G^m(t, \Delta)dt \tag{57}$$

where we have defined

$$G^m(t, \Delta) = \int_0^\infty dz \int_{-z}^0 g_{t,\Delta}^m(\Delta z, z)dz \tag{58}$$

with $g_{t,h}^m(z, y)$ the joint density of $X_m(t)$ and $\dfrac{X_m(t+h) - X_m(t)}{h}$. A change of variable in the integral of Eqn. (57) yields

$$\mathbf{E}[N^m(T_m)] = \lim_{\Delta \rightarrow 0} \int_0^1 T_m G^m(tT_m, \Delta)dt \tag{59}$$

Substituting relation (54) into condition (8) we have that the integrand of Eqn (59) is

uniformly bounded. By Eqn. (55) and dominated convergence, the right hand side of Eqn. (59) converges to the integral

$$\mathbf{E}[N^m(T_m)] \rightarrow \int_0^1 h(t)dt < \infty \tag{60}$$

Thus the assumptions of Proposition 1.1 are satisfied and Eqns. (16) through (27) are valid in the proof of Theorem 1.1. Finally the passage to the limit in Eqn. (30) is justified since the integrand, with respect to r, is bounded above by $G^m(r,\Delta)$.

Combining the above two lemmas gives the following

Proposition 2.1

Fix $\epsilon > 0$ and for every $m \geq 1$ let $X_m(t)$ be a random process which satisfies the conditions in Theorem 2.1. Let $N^m(\sigma,\tau)$ and $N_n^m(\sigma,\tau)$ be the number of upcrossings of zero within $[\sigma,\tau)$ by X_m and the approximation to X_m, ξ_n^m, respectively. Further assume that the rate function of N^m, $h_m(\tau)$, is such that $T_m h_m(\tau T_m)$ converges uniformly to a function $h(\tau)$. Then for $p_m(t)$ the probability that $\tilde{N}^m(0,t) = N^m(0,tT_m)$ exceeds zero and $p^*(t)$ the probability that a Poisson count process with rate function $h(\tau)$ exceeds zero in the interval $[0,t) \in [0,1]$ there exists an $M > 0$ such that

$$\tag{61}$$

$$|p_m(t) - p^*(t)| \leq \lim_{n\to\infty} \sum_{i=0}^{2^n-1} |q^m(t_i)| \exp(-\int_0^t T_m h_m(\tau T_m)d\tau) + \epsilon, \quad m > M$$

where

$$q^m(t_i) = P(B_{t_i,t_{i+1}}^m, \bar{B}_{0,t_i}^m) - P(B_{t_i,t_{i+1}}^m)P(\bar{B}_{0,t_i}^m) \tag{62}$$

As usual B_{t_i,t_j}^m is the event that $N_n^m(t_i,t_j)$ exceeds zero for $t_i, t_j \in \{t_l\}_{l=0}^{2^n}$ which is an increasingly dense partition of the interval

$[0,t)$.

Proof

Let $p_m^{\bullet}(t)$ be the probability that a Poisson count process on $[0, 1]$ with the same rate function as that of $\bar{N}^m$, $T_m\, h_m\, (\tau T_m)$, exceeds zero in $[0,t) \subset [0,1]$. Then we have by the triangle inequality

$$| \; p_m\, (t) - p^{\; \bullet}(t)| \leq | \; p_m\, (t) - p_m^{\bullet}(t)| + | \; p^{\; \bullet}(t) - p_m^{\bullet}(t)| \qquad (63)$$

By Lemma 2.3 X^m satisfies Theorem 1.1 so that by Lemma 2.2 the first magnitude difference on the right of (63) is bounded

$$|p_m\, (t) - q_m\, (t)| \leq \lim_{n \to \infty} \sum_{i=0}^{2^n-1} |q^{\;m}(t_i)| \; exp(-\int_0^t T_m\, h_m\, (\tau T_m\,)d\tau) \quad , \quad t \in [0, 1] \qquad (64)$$

By the finiteness of the integral of h_m for all m, (56) of Lemma 2.3, and the continuity of the exponential function the convergence of $T_m\, h_m\, (\tau T_m\,)$ to $h(\tau)$ implies that the second magnitude difference on the right of the inequality, Eqn. (63), is less than ϵ for sufficiently large m.

Another fact that we will need takes the form of the following generalization of Lemma 2.2.3 in [7].

Lemma 2.4

Let X_m satisfy (6) through (8), be u-a mixing and satisfy (38) and (39). Then given $\epsilon > o$, integers $r > 0$, $K > 0$, positive quantities l, $l = o(T_m / K)$ and t, $\dfrac{T_m}{K} < t < T_m$, we have for m sufficiently large

$$(65)$$

$$P(N_n^m(t-l,t) > 0,\; N_n^m(t-l) = 0) \leq (\; \frac{r}{r+1}\;) \frac{1}{r+1} \; + (2r-1)\alpha_{m,l} \; + \epsilon$$

To prove this we need the following

Lemma 2.5

Let conditions surrounding Eqns. (38) and (39) hold for the upcrossing process $N_n^m(t)$, $t \in I_m$ and assume $\mathbf{E}[N^m(T_m)]$ is finite. Let J be a subinterval of I_m whose length l is $o(T_m)$. Define the collection of r disjoint sets $\{H_i\}_{i=0}^r$ each of length less than the parameter h_o in condition (39) and covering J, $J = \bigcup_{i=1}^r H_i$. Then the following holds

$$\lim_{m,n \to \infty} [\sum_{i=1}^r \mathbf{E}[N_n^m(H_i)] - P(N_n^m(J) > 0)] = 0 \tag{66}$$

Proof

By condition (39) given $\epsilon > 0$ there exists an integer M_1 sufficiently large such that for $m > M_1$

$$\sum_{i=1}^r \mathbf{E}[N_n^m(H_i)] - P(N_n^m(J) > 0) \tag{67}$$

$$\leq \sum_{i=1}^r P(N_n^m(H_i) > 0)(1 + \epsilon) - P(N_n^m(J) > 0)$$

Using the same idea as in the proof of Theorem 1.1 linking upcrossing events with first instances of upcrossing events we can write the second term on the right of Eqn. (67) in a manner analogous to Eqn. (18)

$$P(N_n^m(J) > 0) = \sum_{i=1}^r P(N_n^m(H_i) > 0, N_n^m(\bigcup_{j=1}^{i-1} H_j) = 0) \tag{68}$$

Therefore combining Eqn. (68) and Eqn (67) for $m > M_1$

$$\sum_{i=1}^r \mathbf{E}[N_n^m(H_i)] - P(N_n^m(J) > 0)$$

$$\leq \sum_{i=1}^r P(N_n^m(H_i) > 0, N_n^m(\bigcup_{j=1}^{i-1} H_j) > 0) + \epsilon \sum_{i=1}^r P(N_n^m(H_i) > 0) \tag{69}$$

The first term on the right hand side of Eqn. (69) is bounded by the expression whose limit is zero in Eqn. (38)

$$\sum_{i=1}^{r} P(N_n^m(H_i) > 0, \ N_n^m(\bigcup_{j=1}^{i-1} H_j) > 0)$$

$$\leq \sum_{i=1}^{r} \sum_{t_i \in H_i} P(N_n^m(t_i, t_{i+1}) > 0, \ N_n^m(t_i, t_i - l) > 0) \qquad (70)$$

$$\leq \sum_{t_i = l_m}^{T_m} P(N_n^m(t_i, t_{i+1}) > 0, \ N_n^m(t_i, t_i - l) > 0)$$

Likewise by the "rarefaction" condition there exists an M_2 such that the last term in the inequality (70) is less than ϵ for $m > M_2$. Therefore for $m > \max(M_1, M_2)$

$$\sum_{i=1}^{r} \mathbf{E}[N_n^m(H_i)] - P(N_n^m(J) > 0) \leq \epsilon(1 + \sum_{i=1}^{r} P(N_n^m(H_i) > 0)) \qquad (71)$$

Since ϵ can be arbitrarily small, as we let m go to infinity the limiting result, Eqn. (66), follows from Eqn. (71).

By the disjointness of the H_j in Lemma 2.5 we can assert that, for any interval J with length of asymptotic order T_m, if the assumptions of the lemma are satisfied then

$$\lim_{m \to \infty} |\mathbf{E}[N_n^m(J)] - P(N_n^m(J) > 0)| \to 0 \ , \quad as \ m \to \infty \qquad (72)$$

With the above we can now show Lemma 2.4.

Proof of Lemma 2.4

By the previous lemma and its implication (72), for $t \in [l, T_m)$: given $\epsilon > 0$ for large enough m

$$|P(N_n^m(t-l, t) > 0) - \mathbf{E}[N_n^m(t-l, t)]| < \epsilon \qquad (73)$$

Since $l = o(\frac{T_m}{K})$, for sufficiently large m the Uniform Denseness Condition implies that, for some $r > 0$, there are r intervals, $\{E_i\}_{i=1}^{r} \subset [0, \frac{T_m}{K} - l)$, each separated by l such that

$$|\mathbf{E}[N_n^m(E_i)] - \mathbf{E}[N_n^m(t-l,t)]| < \epsilon , \quad i = 1, 2, \ldots, r \tag{74}$$

Furthermore r can be taken as arbitrarily large as $m \to \infty$. Therefore combining (74), (73), and (72) with $J = E_i$, we have for sufficiently large m

$$|P(B_{E_i}^m) - q^{(m)}| < 3\epsilon \tag{75}$$

where the triangle inequality has been invoked and we have defined

$$q^{(m)} = P(B_{t-l,t}^m) \tag{76}$$

and as before B_J^m is the event that $N_n^m(J)$ exceeds zero. Here we will drop the index m for convenience. The left hand side of inequality (65) can thus be bounded

$$P(B_{t-l,t}, \bar{B}_{0,t-l}) \leq P(B_{t-l,t}, \bigcap_{i=1}^{r} \bar{B}_{E_i}) \tag{77}$$

or

$$P(B_{t-l,t}, \bar{B}_{0,t-l}) \leq P(\bigcap_{i=1}^{r} \bar{B}_{E_i}) - P(\bar{B}_{t-l,t}, \bigcap_{i=1}^{r} \bar{B}_{E_i}) \tag{78}$$

Now from Lemma 2.1

$$|P(\bigcap_{i=1}^{r} \bar{B}_{E_i}) - \prod_{i=1}^{r} P(\bar{B}_{E_i})| \leq (r-1)\alpha_{m,l} \tag{79}$$

and

$$|P(\bar{B}_{t-l,t}, \bigcap_{i=1}^{r} B_{E_i}) - P(\bar{B}_{t-l,t}) \prod_{i=1}^{r} P(\bar{B}_{E_i})| \leq r\alpha_{m,l} \tag{80}$$

Hence the right hand side of the inequality (78) does not exceed

$$(1-q)^r + O(\epsilon) + (r-1)\alpha_l + (1-q)^{r+1} + O(\epsilon) = r\alpha_{m,l} \tag{81}$$

$$= q(1-q)^r + (2r-1)\alpha_{m,l} + O(\epsilon) \tag{82}$$

where $O(\epsilon)$ goes to zero as ϵ goes to zero. The first term in (82) has a maximum value of $(\frac{r}{r+1})^r \frac{1}{r+1}$ which is verified by investigating derivatives in the unit interval. We lose no generality by replacing $O(\epsilon)$ by ϵ in (82) which yields Lemma 2.4.

The major technical portion of the proof of Theorem 2.1 is contained in the following proposition.

Proposition 2.3

If the a.s. continuous processes $X_m(t)$ satisfy the conditions stated in the premise of Theorem 2.1 then for any semiclosed interval I contained in $[0,1]$

$$P(\tilde{N}^m(I) > 0) \to 1 - exp(-\int_I h(\tau)d\tau) \qquad (83)$$

Proof

First fix I greater than zero and K a positive integer. We reproduce Eqn. (62) here for clarity. As in Eqn. (12) of Theorem 1.1, for the sequence X_m, $m = 0, 1, \ldots$ we have the quantities $q^m(t_i)$ on the 2^n point grid $\{t_i\}_{i=0}^{2^n}$

$$q^m(t_i) = P(B_{t_i, t_{i+1}}^m, \bar{B}_{0,t_i}^m) - P(B_{t_i, t_{i+1}}^m)P(\bar{B}_{0,t_i}^m) \qquad (84)$$

Partition the interval $[0, T_m)$ into K parts so that the sum on the right of Eqn. (61) of Proposition 2.1 can be represented as

$$\sum_{i=0}^{2^n - 1} |q^m(t_i)| = \sum_1 |q^m(t_i)| + \ldots + \sum_K |q^m(t_i)| \qquad (85)$$

where $\sum_k$ denotes summation over the intersection of the grid $\{t_i\}_{i=0}^{2^n}$ and the k-th partition element, $k = 1, 2, \ldots, K$.

Fix $\epsilon > 0$ and let m be sufficiently large so that Lemma 2.4 holds. Consider the final K-1 terms in (85)

$$\sum_2 |q^m(t_i)| + \ldots + \sum_K |q^m(t_i)| \qquad (86)$$

Now for each $q^m(t_i)$ in (86) we add and subtract terms so as to isolate the mixing dominated quantities of the form (36). That is we obtain via the triangle inequality

$$|q(t_i)| \leq |P(B_{t_i,t_{i+1}}, \bar{B}_{0,t_i-l}) - P(B_{t_i,t_{i+1}})P(\bar{B}_{0,t_i-l})| \tag{87}$$
$$+ |P(B_{t_i,t_{i+1}})P(\bar{B}_{0,t_i-l}) - P(B_{t_i,t_{i+1}})P(\bar{B}_{0,t_i})|$$
$$+ |P(B_{t_i,t_{i+1}}, \bar{B}_{0,t_i}) - P(B_{t_i,t_{i+1}}, \bar{B}_{0,t_i-l})|$$

Where we have suppressed dependencies on m for notational simplicity. Using the mixing condition (36) on the first term to the right of the inequality (87) and using simple set identities for the other two terms we have

$$|q(t_i)| \leq P(B_{t_i,t_{i+1}}) \left[\alpha_{m,l} + P(B_{t_i-l,t_i}, \bar{B}_{0,t_i-l}) \right] + P(B_{t_i,t_{i+1}}, B_{t_i-l,t_i}, \bar{B}_{0,t_i-l}) \tag{88}$$

Finally applying Lemma 2.4 to the second term in brackets [ ] in (88) and noting that by monotonicity the third term in (88) is bounded

$$P(B^m_{t_i,t_{i+1}}, B^m_{t_i-l,t_i}, \bar{B}^m_{0,t_i-l}) \leq P(N^m_n(t_i, t_{i+1}) > 0, N^m_n(t_i-l, t_i) > 0) \tag{89}$$

we get by substituting the inequality (88) in (85)

$$\sum_{i=0}^{2^n-1} |q^m(t_i)| \tag{90}$$

$$\leq \sum_1 |q^m(t_i)| + \left[2r\alpha_l + \left(\frac{r}{r+1} \right)^r \frac{1}{r+1} + \epsilon \right] \sum_{i=\varsigma}^{2^n-1} P(B^m_{t_i,t_{i+1}})$$

$$+ \sum_{i=\varsigma}^{2^n-1} P(N^m_n(t_i,t_{i+1}) > 0, N^m_n(t_i-l,t_i) > 0)$$

here $\varsigma \in \{t_i\}_{i=0}^{2^n}$ is the rightmost point contained in the first partition element, $[0, \frac{T_m}{K}]$, of the K-th order partition. Now applying the relation (31) and Lemma 1.2 to the first term to the right of the inequality (90) for n sufficiently large

$$\sum_1 |q^m(t_i)| \leq \sum_1 P(B^m_{t_i,t_{i+1}}) \leq E[N^m_{n_1}] + \epsilon \tag{91}$$

Where $N^m_{n_k}$ is as defined in Eq. (41). Likewise

$$\sum_{i=\varsigma}^{2^n-1} P(B^m_{t_i,t_{i+1}}) \leq \sum_{i=0}^{2^n-1} P(B_{t_i,t_{i+1}}) \leq E[N^m(T_m)] + \epsilon \tag{92}$$

which gives via Eqn. (90)

$$\lim_{n \to \infty} \sum_{i=0}^{2^a-1} |q^m(t_i)|$$

$$\leq \lim_{n \to \infty} E[N_{n_1}^m] + [2r\alpha_{m,l} + (\frac{r}{r+1})^r \frac{1}{r+1} + \epsilon]\{E[N^m(T_m)]\} \qquad (93)$$

$$+ \lim_{n \to \infty} \sum_{i=\varsigma}^{2^a-1} P(N_n^m(t_i, t_{i+1}) > 0, N_n^m(t_i - l, t_i) > 0)$$

Therefore taking the limit as $m, l \to \infty$, $l = o(T_m)$, the first term to the right of (93) goes to zero by Asymptotic Uniform Negligibility, (41), and the finiteness of the limit $\lim_{m \to \infty} E[N^m(T_m)] = E[\tilde{N}(1)]$. The second term converges to a quantity not exceeding $[\frac{1}{r} + \epsilon] E[\tilde{N}(1)]$. However as m becomes unbounded r can be made arbitrarily large and ϵ can be made arbitrarily small, by Lemma 2.4, thus the second term is negligible. Finally the rarefaction condition, Eqn. (38), asserts that the third term vanishes.

Hence by Proposition 2.1, for $I = [0, t]$

$$p_m(T_m t) \to p^*(t) = 1 - \exp(-\int_0^t h(\tau)d\tau), \quad t \in [0,1) \qquad (94)$$

Proposition 2.3 asserts that the probability that the normalized upcrossing process $\tilde{N}^m$ is greater than zero in any interval contained in [0,1) is the same as the corresponding probability for a Poisson counting process N^* in the limit as $m \to \infty$. To show the stronger result that $\tilde{N}^m$ actually converges in distribution to a Poisson process we will follow Leadbetter [7] in making use of a theorem in [5]. Using the nomenclature in [5] a point process N is *a–regular* if for every collection of intervals I contained in $\Upsilon_{[0,1]}$, the Borel sets on [0,1], there exists some array $\{I_{mk}\} \subset \Upsilon_{[0,1]}$ of finite covers of I (one for each $m = 1,2,...$) such that

$$\lim_{n \to \infty} \limsup_{m \to \infty} \sum_k P(N^m(I_{nk}) \geq a) = 0 \qquad (95)$$

We state the following special case of Theorem 4.7 in [5].

Lemma 2.6

Let $\tilde{N}^m$ be a sequence of point processes and N^* a Poisson process both defined on $[0,1)$. Then $\tilde{N}^m$ converges in distribution to N^* if and only if $\tilde{N}^m$ is 2-regular and

$$\lim_{m \to \infty} P(\tilde{N}^m(U) = 0) = P(N^*(U) = 0) \tag{96}$$

for all U of the form

$$U = \bigcup_{k=1}^{r} \Upsilon_i \quad , \quad \Upsilon_i \subset \Upsilon_{[0,1]}$$

for $r > 1$, and

$$limsup_{m \to \infty} \mathbf{E}[\tilde{N}^m(I)] \leq \mathbf{E}[N^*(I)] < \infty \tag{97}$$

for $I \subset \Upsilon_{[0,1]}$

We now proceed to the proof of Theorem 2.1 which at this point only involves showing that $\tilde{N}^m$ of the theorem satisfies the conditions in Lemma 2.6.

Proof of Theorem 2.1

Without loss of generality we assume that the collection of intervals I in the *a–regularity* condition and in (97), and the Υ in (96) are sets of disjoint intervals. For each m, $m = 1, \dots$, define the increasing set of disjoint covers of I: $\{J_{mk}\}$, $k = 1, 2, \dots r_m$, with each J_{mk} of length l_m / T_m (recall $l_m = o(T_m)$). Assume for definiteness that J_{mk} are ordered such that the left endpoints are strictly increasing as k increases. With N^m as in Theorem 2.1 and $\tilde{N}^m(\tau) = N^m(\tau T_m)$ we have

$$\sum_{k=1}^{r_m} P(\tilde{N}^m(J_{mk}) > 1) = \sum_{k=2}^{r_m} P(\tilde{N}^m(J_{mk}) > 1) + P(\tilde{N}^m(J_{m1}) > 1) \tag{98}$$

$$\leq \lim_{n \to \infty} \sum_{k=1}^{r_m} \sum_{t_i \in J_{km}} P(N_n^m(t_i^k, t_{i+1}^k) > 0, N_n^m(t_i^k - l_m, t_i^k) > 0) + \mathbf{E}[\tilde{N}^m(J_{m1})]$$

Where $\{t_i^k\}_{i=1}^{n^k}$ are increasingly dense partitions of J_{mk}, for $k = 1, \dots r_m$ respectively. The first term on the right of the inequality (98) is bounded by the expression in the Rarefaction condition, Eqn. (38) while the second term converges to zero by Asymptotic

Uniform Negligibility, (41), and the finiteness of $E[\tilde{N}^m]$ (Lemma 2.3). Taking the limit of Eqn. (98) as $m \rightarrow \infty$ we have that $\tilde{N}^m$ is 2-*regular* in the sense of (95).

Fix $r > 0$. Because of the absolute continuity of the distributions of N^m the intervals Υ_i in (96) can be taken as having no common boundary points. Therefore by mixing, Lemma 2.1, for any collection of disjoint intervals $\Upsilon_1, \Upsilon_2, \ldots, \Upsilon_r$ contained in $[0,1]$

$$|P(\bigcap_{i=1}^{r} N^m(T_m \Upsilon_i) > 0) - \prod_{i=1}^{r} P(N^m(T_m \Upsilon_i) > 0)| \rightarrow 0 , \quad as \ m \rightarrow \infty \qquad (99)$$

where we have adopted the operator notation for T_m: $T_m[\sigma,\nu) = [T_m \sigma, T_m \nu)$ for $0 \leq \sigma < \nu < 1$. Eqn. (99) and Proposition 2.3 thus imply that

$$P(\bigcap_{i=1}^{r} N^m(T_m \Upsilon_i) = 0) \rightarrow \exp(-\sum_{i=1}^{r} \int_{\Upsilon_i} h(\tau)d\tau) \qquad (100)$$

We have from the additivity of the expectation operator over disjoint intervals and Lemma 2.3 applied to general intervals contained in $[0, T_m)$

$$E[\tilde{N}^m(I)] \rightarrow E[N^*(I)] < \infty \qquad (101)$$

Therefore the assumptions stated in Lemma 2.6 are valid for $\tilde{N}^m$ and Proposition 2.3 establishes Theorem 2.1.

IV. Discussion

In this chapter two results were derived in the context of level crossing probabilities. First, a representation of the probability of getting one or more upcrossings in an interval by a general random process was presented. This representation in effect isolates the portion of the upcrossing probability due to the intensity function of the upcrossings, from a correction term, which characterizes the deviation of the upcrossing probability from an associated inhomogeneous Poisson probability. The correction term depends on the degree to which

the upcrossings can be modeled as an independent increment process. By identifying conditions which asymptotically force the correction term to zero a second result was made possible: that a certain time normalized version of the upcrossing process can be made to converge in distribution to the inhomogeneous Poisson law.

Of immediate interest to us is the use of an inhomogeneous Poisson model for the probability of a random process upcrossing a curve or level. Future investigations of the specific form of the correction term, associated with the probability representation of Thm. 1.1, could lead to useful expressions for the approximation error incurred by using such simple first moment approximations. For the asymptotic result, Thm. 2.1, the fairly complicated conditions, rarefaction, uniform denseness and asymptotic uniform negligibility, could use some simplification. In particular, rarefaction and uniform denseness could be replaced by conditions involving probability statements about the maximum process over the interval I, $\max_{\tau \in I} X(\tau)$, analogous to [7]. For specific probability models of the random process $X(t)$ of interest, e.g. Gauss-Markov or Rayleigh as in [7], one would expect the replacement conditions to be more easily verified, than the conditions used here.

REFERENCES

Chapter 4

1. Ylvisaker, N.D., "The expected number of zeros of a stationary Gaussian Process", *Ann. of Math. Statist.*, vol 36 pp 1043-1046, 1965.

2. Leadbetter, M.R., "On Crossings of Levels and Curves by a Wide Class of Stochastic Processes", *Ann. of Math. Statist.* vol 37, pp. 260-267, 1966.

3. Rosenblatt, M., Random Processes, N.Y., Oxford University Press, 1962.

4. Vidyasagar, M., Nonlinear Systems Analysis, Englewood Cliffs, Prentice-Hall, 1978.

5. Kallenberg, O., Random Measures, Berlin, Akademie-Verlag, 1975.

6. M. Kac and D. Slepian, "Large excursions of Gaussian processes," *Ann. of Math. Statist.* vol. 18, pp. 383-397, 1947b.

7. Leadbetter, M.R., "Extreme value theory for continuous parameter stationary processes", *Zeitschrift fur Wahrscheinlichkeitstheorie und Verwandte Gebiete,* vol. 60, pp. 1-20, 1982.

8. I. F. Blake and W. C. Lindsay, "Level-crossing problems for random processes," *IEEE Trans. on Inform. Theory,* Vol. IT-19, pp. 295-315, 1973.

9. V. A. Volkonskii and Y. A. Rozanov, "*Some limit theorems for random functions. II,*" Theory of Prob. and its App. (English Translation), Vol. 6, No. 2, pp. 186-198, 1961.

10. M. R. Leadbetter, "*Point processes generated by level crossings,*" P. A. W. Lewis, ed., Stochastic Point Processes, New York, Wiley, 1972.

11. H. Cramer, *Mathematical Methods of Statistics,* Princeton, N.J.: Princeton University, 1951.

Chapter 5
A Level-Crossing Approach to Large
Deviations in Time Delay Estimates

I. Introduction

In this chapter the performance of the GCC is investigated for low observation-time--signal-bandwidth (BT) products. The performance of the GCC has been characterized by the estimator variance obtained by a linearization technique first introduced by P. Schultheiss [2]. The resulting variance expression, in [2], is equivalent to the Cramer-Rao lower bound (CRLB), but it is only valid for "good" estimates, i.e. those close to the true time delay. For lowpass signal spectra, the estimate will be good in the high signal-to-noise ratio (SNR) case; however, in low SNR situations, the estimates can be far removed from this local region. In the latter case, the local variance can underestimate the actual variance of the estimate by orders of magnitude. For narrowband signal spectra, the GCC may never achieve the CRLB [19]. The disagreement between local and actual variance was shown analytically to be significant in the narrowband signal case by Chow and Schulthiess, using a form of the Barankin bound [4]. The deterioration in performance beyond that predicted by the local variance occurs in what Chow and Schulthiess call the "ambiguity dominated region" on the SNR versus bandwidth plane. This region is delineated by a SNR threshold beyond which spurious peaks give rise to exceedingly poor estimates of time delay. In [18], Weiss and Weinstein constructed a lower bound on the error variance based on the Ziv-Zakai bound. The resultant bound is tighter than the Barankin bound of [4] and reveals the presence of two seperate SNR thresholds where the lower bound on the error begins successive rapid increases in magnitude.

In this chapter, we develop approximations to the probability of large error and the global variance of a given GCC implementation. The approximations presented here are sensitive to large errors and can be applied to arbitrary signal and noise spectra.

In a recent paper [3], Ianniello investigated the effect of peak anomaly on the performance of the simple cross-correlator for broadband lowpass signals spectra. He obtained an approximate expression for the probability of anomaly using a method analogous to the approximation of the probability of error in Pulse Position Modulation communication systems. A model for the global variance was proposed, therein referred to as the Correlator Performance Estimate (CPE), which took the anomalous estimates of delay as lying uniformly throughout the *a priori* region of delay. The above approximations are quite accurate for the cases studied in the references [3],[14], where the cross-correlation function has no sidelobes over the majority of the *a priori* region of delay. However, for the general case large positive sidelobes can introduce errors not accounted for in Ianniello's approximation.

For bandpass signals, Ianniello made a distinction between anomaly and peak ambiguity in [19]. Essentially, ambiguity denotes the occurrence of spurious peaks near the high frequency sidelobes of the cross-correlation, within the central lobe of the envelope of the correlation function. Anomaly, on the other hand, is error which specifically occurs outside of this central lobe. To the best of our knowledge, no unified treatment of both kinds of large error has been undertaken.

Here a different approach to error modeling is investigated, using certain level crossing statistics of the GCC. This approach will be sufficiently general to handle large error for small time-bandwidth products, and for high frequency bandpass signals, since it incorporates the time varying statistics of the GCC tra-

jectory into the error analysis. Since no distinction need be made between anomaly and ambiguity in this approach, the all encompassing term "peak ambiguity" will be used to denote large error. In what follows, the approximate equivalence of peak ambiguities with suitably defined level crossings by the GCC trajectory will be demonstrated. The probability of peak ambiguity can then be represented in terms of an infinite series involving the moments of the number of these level crossings over the *a priori* region. This expression is exact for arbitrary signals as long as these moments are finite and smooth. In particular, the time varying statistics of the GCC output (i.e. sidelobes) are taken into account by the level crossing model. Here the emphasis is on predicting the large error performance using knowledge of the first moment only, that is the average number of peak ambiguities over the *a priori* interval. Based on the first moment, a general Poisson model for peak ambiguity is investigated, motivated by asymptotic considerations studied in the previous chapter. The Poisson model leads to an expression for the global variance which incorporates the variance approximation of [2] and the intensity function of the ambiguous peaks, the instantaneous average number of peak ambiguities. The resultant expression for the global variance, $var[\hat{D}]$, has the form

$$var[\hat{D}] = P_e \, var_A \, + (1 - P_e)var_L$$

where: P_e is the probability of peak ambiguity, var_L is a truncated version of the local variance [2], and var_A is a measure of the mean-square distance of the locations of the ambiguous peaks from the true time delay (the second moment of the intensity function over time). Specific results for the intensity function of the ambiguities are then derived using a Gaussian approximation to the GCC output and an asymptotic analysis is performed for broadband signal spectra and large time-bandwidth products (length of *a priori* region times bandwidth ($D_m \, T$)). One result of the analysis is that at high SNR the locations of ambiguities,

specifically the intensity function, become essentially independent of time, while at low SNR the time dependency behaves approximately as $1/(B\tau)^2$, where B is the signal bandwidth and τ ranges over the *a priori* region. The limiting form, as $B\tau \rightarrow \infty$, of the average number of ambiguities are valid for all ranges of SNR and are simple to evaluate, involving no integrations.

For the case of signals with flat lowpass spectra, it can be shown that the Poisson approximation to the large error performance of the GCC is conservative; it over-estimates the actual probability of ambiguity. However, the threshold SNR, delimiting the onset of severe peak ambiguity, is underestimated on the order of less than 3dB for the cases of lowpass signal spectra considered here. Motivated by heuristics, a model for the approximation error between the Poisson and the actual ambiguity probability is discussed. Using results obtained in Chapter 4, an improved Poisson approximation is derived which adjusts the intensity function of ambiguities downward. The Poisson motivated approximations to ambiguity probability and global variance are then compared to Ianniello's probability of anomaly and the CPE for the broadband signal cases studied in [3],[14].

Finally, a numerical study of the behavior of the average number of ambiguities and the Poisson probability and variance approximations is undertaken for narrowband bandpass signal spectra. For low BT, a positive indication of the location of the lower of the SNR thresholds, and the SNR value indicating the onset of the domination of *a priori*-information over the observations, given by the Ziv-Zakai Lower Bound [18], is provided by the approximations. However, the approximations indicate that, for low BT, the upper SNR threshold, indicated in [18], is overly optimistic, i.e. the CRLB is unattainable using the GCC. Therefore, in this case the GCC is not an efficient (optimal) estimator. For high BT, the location of the higher SNR threshold, where the attainable variance begins to

deviate from the CRLB, is positively identified by the Poisson approximations. This implies that the GCC is an optimal estimator for high BT. The above indications are in agreement with conclusions drawn from a local (small error) analysis of estimator variance, undertaken by Ianniello et. al. for bandpass signals [19].

II. Problem Statement

We briefly recall the setting of the two sensor passive time delay estimation problem defined in Chapters 2 and 3. Available are the outputs of two sensors, $x_1(t)$ and $x_2(t)$, over an interval of time $t \in [0, T]$. $x_1(t)$ contains a signal, $s(t)$, while $x_2(t)$ contains a delayed version $s(t-D_o)$. Both signal components are observed in additive mutually uncorrelated broadband noises, $n_1(t)$ and $n_2(t)$. We assume that $s(t)$, $n_1(t)$, and $n_2(t)$ are ergodic, zero mean, stationary Gaussian random processes possessing spectral densities $G_{ss}(\omega), G_{n1}(\omega)$ and $G_{n2}(\omega)$ respectively. In addition $D_o \in [-D_m, D_m]$, some *a priori* region of delay. The signal auto-correlation, $R_{ss}(\tau) = E[s(t)s(t-\tau)]$, is assumed to be essentially zero for $|\tau| > T_c$, the correlation time of the signal, which is also the independence time for the present case. This implies that the auto and cross-correlation functions of the observations, $R_{ij}(\tau) = E[x_i(t)x_j(t-\tau)]$, are non-zero over at most a distance T_c from the time origin.

The cross-correlation estimate of time delay, $\hat{D}$, is the location in time, within $[-D_m, D_m]$, at which the absolute maximum of the (coherent or incoherent) sample cross-correlation function occurs. By ergodicity the sample cross-correlation, $\hat{R}_{12}(\tau)$, converges to the true cross-correlation, $R_{12}(\tau) = R_{ss}(\tau-D_o)$, in the limit of large observation time, T. For finite T, however, $\hat{R}_{12}(\tau)$ may not have its absolute maximum at D_o. Recall, from Chapter 2 Eq. (5), the sample cross-correlation can be expressed as the sum of

signal type terms and noise type terms

$$\hat{R}_{12}(\tau) = \hat{R}_{ss}(\tau - D_o) + \hat{R}_n(\tau) \qquad (1)$$

The first term in (1) is uniquely dependent on the signal and has an absolute maximum at D_o. The second term represents a zero-mean noise in what can be called the cross-spectral domain in contrast to the domain of the observations, Eq. (1) of Chapter 1. This term, $\hat{R}_n(\tau)$, masks the global peak at D_o of $\hat{R}_{ss}(\tau - D_o)$ and thus degrades the accuracy of the estimator. For non-white signals one can do better by performing Generalized Cross-Correlation obtained by filtering the sample cross-correlation, with a filter with transform $W(w)$, to resolve the peak at $\tau = D_o$. For representatives of the GCC class see [1] or Chapter 2. In the following, only the simple cross-correlation estimate will be studied, that is $W(w) = 1$ in the GCC. Other GCC's can be handled by dealing with the $W(w)$-derived signal and noises, defined by the equivalent scenario where simple cross-correlation of the $W(w)$-derived observations yields the identical GCC output as for the original observations (Substitute the filtered process $s(t) * w^+(t)$ of Chapter 3, Eq. (3), for $s(t)$ in the observation model of Chapter 2, Eq. (1)).

In general, the peak location algorithm involves quantization of R^g and a global search over $[-D_m, D_m]$. Thus a resolution parameter, ϵ, related to the granularity of the quantization, can be associated with the peak detection operation. This will introduce additional errors into the cross-correlation estimate due to quantization error.

The local variance of the GCC is derived in [2] as a result of linearizing the first derivative of $E[R^g(\tau)]$ in the neighborhood of D_o. Consequently the analysis is exact as long as $R_{ss}(\tau)$ does not deviate from an inverted parabola centered about zero with maximum $R_{ss}(0)$ and curvature: $\dfrac{d^2 R_{ss}(\tau)}{d\tau^2}\Big|_{\tau=0}$. For

the signals considered in this chapter the above consideration implies that the region of linearity for the local variance analysis lies approximately between the pair of zeros of $R_{ss}(D_o)$ which are closest to D_o (See Fig. 2.2). For a spectrally broadband signal $s(t)$, with single-sided bandwidth B these zeros occur at $D_o - \frac{2\pi}{B}$ and $D_o + \frac{2\pi}{B}$. For a bandpass signal, at center frequency w_s the region of accuracy of the local variance is delimited by $D_o - \frac{\pi}{2w_s}$ and $D_o - \frac{\pi}{2w_s}$ (See Fig. 2.2a and Fig. 2.2b for an example of a bandpass signal and its envelope). Therefore, if the probability that the error falls outside the linear region is high, the local variance will give an overly optimistic indication of the performance of the GCC. This probability will be the most critical in the bandpass case because of the high level sidelobes of R^g (Fig. 2.2a). In the sequel we will refer to the "linear region", mentioned above, as the "small error region", and the remainder of the *a priori* interval as the "large error region".

Ianniello considered an approximation to the probability of anomaly, for broadband lowpass signals, by looking at the probability that a number of test points on the correlation trajectory in the large error region exceed the value of the trajectory at the true delay D_o [3]. To make the computation of the above probability tractable, the maximum number of independent samples of the correlator output were chosen as test points and these were assumed to be Gaussian distributed. In our notation there are approximately $M = 2D_m/T_c$ of these points (including the sample at D_o) and the approximation takes the form

$$P_e = 1 - \int_{-\infty}^{\infty} \phi(x - \alpha)\left\{ \Phi(\beta x)\right\}^{M-1} dx \qquad (2)$$

where α and β are constants involving SNR and the time - bandwidth product BT. ϕ is the standard univariate Gaussian density:

$$\phi(x) = \frac{1}{\sqrt{2\pi}} \; e^{-\frac{x^2}{2}} \tag{3}$$

and Φ is the Gaussian distribution function:

$$\Phi(x) = \int\limits_{-\infty}^{\infty} \phi(s)ds \tag{4}$$

Equation (2) is only applicable when the sidelobe activity of the correlation function is neglected and, in general, it must be numerically evaluated. The Correlator Performance Estimate (CPE) proposed by Ianniello assumes that the peak ambiguities are uniformly distributed over the *a priori* interval. The resultant expression for the CPE involves the approximation (2) for P_e

$$CPE = P_e \; \frac{D_m^2}{3} \; + (1 - P_e)var_L \tag{5}$$

where var_L is the linear approximation to the variance derived by Schultheiss [2].

In the next section the probability of large error will be approximated in terms of level crossing probabilities. Then in Section V a variance approximation is developed which can be applied to situations where the locations of ambiguities are distinctly not uniform over the *a priori* interval. These approximations will be more general than the approximations (2) and (5), since they will be able to deal with both broadband and narrowband signals which may display large sidelobes.

III. A Level Crossing Interpretation For Large Error

In this section a model for the occurence of large errors, or peak ambiguity, is developed. By definition a peak ambiguity occurs when the global peak detection scheme indicates that the absolute maximum of the output of the GCC, $R^g(\tau)$, resides outside of the small error region $[D_o-\delta, \; D_o+\delta]$, e.g. $\delta = \frac{2\pi}{B}$ or $\delta = \frac{\pi}{2w_s}$ for lowpass or bandpass signal spectra of single sided bandwidth, $B/2$,

and center frequency, w_s, respectively.

As an approximation we will characterize the peak ambiguity by the exceedance of $R^g(D_o)$ by $R^g(\tau)$ outside of the local region. This essentially considers $R^g(D_o)$ as a representative of the entire correlation trajectory within the local region, an approximation implicit in Ianniello's estimate discussed in the previous section. At worst, the effect of the approximation will be to overestimate the probability of peak ambiguity. More precisely let ϵ be the minimum magnitude resolution of the peak detection algorithm. Then a large error is characterized by the event

$$R^g(\tau) - R^g(D_o) > -\epsilon \ , \quad some \ \tau \in \mathbf{H}_\delta \qquad (6)$$
$$\mathbf{H}_\delta = \{t : t \in [-D_m, D_o - \delta] \cup [D_o + \delta, D_m]\}$$

For convenience define the ambiguity process $\Delta R^g(\tau) = R^g(\tau) - R^g(D_o)$. In what follows we will assume a.s. continuity of the ambiguity process in its time argument. This will permit probability statements concerning the set of time samples $\{R^g(t_i)\}$, for a dense set of points t_i, to carry over to the entire continuous trajectory of R^g. Define the occurence of peak ambiguities over some set of intervals Θ contained in $\mathbf{H}_\delta$ by $A_\Theta = \{\Delta R^g(\tau) > -\epsilon \ , \quad some \ \tau \in \Theta\}$ and designate the left and right intervals for $\delta > 0$

$$\mathbf{I}_L = [-D_m, D_o - \delta] \qquad (7)$$
$$\mathbf{I}_R = [D_o + \delta, D_m]$$

Then the equation governing the probability of getting a peak ambiguity over $\mathbf{H}_\delta = \mathbf{I}_L \cup \mathbf{I}_R$, P_e, can be written

$$P_e = P(A_{\mathbf{I}_L \cup \mathbf{I}_R}) = P(A_{\mathbf{I}_L}) + P(A_{\mathbf{I}_R}) - P(A_{\mathbf{I}_L} \cap A_{\mathbf{I}_R}) \qquad (8)$$

Eq. (8) can be simplified considerably under the assumptions of very narrowband or very broadband signal spectra. The following comments are considered in detail in the appendices (Appendices A, B and C). For sufficiently

broadband signals, and $\delta = \dfrac{2\pi}{B}$, δ is approximately the independence time, T_c , of $s(t)$. In this case the occurence of level crossings over I_L is approximately independent of occurences in I_R , and Eq. (8) can be expressed as

$$P_e = P(A_{\mathrm{I}_L}) + P(A_{\mathrm{I}_R}) - P(A_{\mathrm{I}_L})P(A_{\mathrm{I}_R}) \tag{9}$$

For the signals and noises of interest the finite dimensional distributions of R^g , i.e. the distributions of $\{R^g(t_i)\}_{i=1}^N$ for $N = 1, 2, \ldots$, depend only on the differences $|t_i - D_o|$. In Appendix B it is shown that as a result of this shift invariance, the probability of peak ambiguity is not dependent on the actual value of D_o if D_m is large with respect to T_c . Without any loss in generality we will take $D_o = 0$. Then I_L and I_R are of equal length and $P(A_{\mathrm{I}_L})$ is equal to $P(A_{\mathrm{I}_R})$, due to the symmetry of the statistics of $R^g(\tau)$ about $\tau = 0$. In this case Eq. (9) becomes

$$P_e = P(A_{\mathrm{I}_L})[2 - P(A_{\mathrm{I}_L})] , \quad \mathrm{I}_L = [-D_m , \dfrac{2\pi}{B}] \tag{10}$$

For narrowband signals it is shown in the appendices that, for spectra of interest, a positive correlation exists between $R^g(\tau)$ and $R^g(-\tau)$ when τ is in the neighborhood of an ambiguity prone sidelobe of R_{ss} . Therefore it is reasonable to make the approximation

$$P(A_{\mathrm{I}_L} \cap A_{\mathrm{I}_R}) = P(A_{\mathrm{I}_L}) \tag{11}$$

That is a peak ambiguity in I_L implies, with high probability, that one occurs in I_R . Hence for narrowband signals we will use

$$P_e = P(A_{\mathrm{I}_L}) \tag{12}$$

where for the lowpass case, $\mathrm{I}_L = [-D_m , -\dfrac{2\pi}{B}]$, and for the bandpass case, $\mathrm{I}_L = [-D_m , -\dfrac{\pi}{2w_s}]$. The representations (10) and (12) allows us to restrict

attention to the single interval I_L, which will be done in the sequel.

If we assign a direction of pursuit to the trajectory $\Delta R^g(\tau)$ over $I = I_L$ then an equivalence can be drawn between peak ambiguities and level crossings of ϵ by ΔR^g. We will let $\Delta R^g(\tau)$ trace out a path as its time argument increases from $-D_m$ to $-\delta$. Define the cumulative distribution of $\Delta R^g(\tau)$ at $\tau = t$

$$F(a;t) = P(\Delta R^g(t) \leq a) \tag{13}$$

Then the probability that an ambiguity occurs in I can be written

$$P(A_I) = P(A_I | \Delta R^g(-D_m) \leq -\epsilon) F(-\epsilon;-D_m) \tag{14}$$
$$+ P(A_I | \Delta R^g(-D_m) > -\epsilon)(1 - F(-\epsilon;-D_m))$$

Where $P(A|B)$ denotes the conditional probability of event A given event B.

A simplification of Eq. (14) is immediate. The conditional probability $P(A_I | \Delta R^g(-D_m) > -\epsilon)$ is equal to one by definition. Rewrite Eq. (14) with the above fact

$$P(A_I) = P(A_I | \Delta R^g(-D_m) \leq -\epsilon) F(-\epsilon;-D_m) + 1 - F(-\epsilon;-D_m) \tag{15}$$

Fix $\delta > 0$, and define the upcrossing count process associated with ΔR^g

$$N(J) = \left\{ \begin{array}{l} \textit{number of times } \Delta R^g \textit{ upcrosses} \\ -\epsilon \textit{ in an interval } J \in H_\delta \end{array} \right. \tag{16}$$

Given ΔR^g started its trajectory below $-\epsilon$, $N(I)$ must be greater than zero for an ambiguity to have occured in I. Since the converse is also true we can write

$$P(A_I) = P_c(N(I) > 0) \, F(-\epsilon;-D_m) + 1 - F(-\epsilon;-D_m) \tag{17}$$

where $P_c(B) = P(B \mid \Delta R^g(-D_m) \leq -\epsilon)$.

Using Eq. (17) in Eq. (10) or Eq. (12) gives an explicit relation between upcrossing probabilities and peak ambiguity probabilities. Thus the determination of the probability of large error is reduced to solving for a level crossing probability. Unfortunately the distribution of level crossings is unknown explicitly

except for certain classes of Markov and pseudo-Markov processes [11]. Motivated by the results on the asymptotic distributions of level crossing probabilities in Chapter 4, an approximation to $P_c(N(I) > 0)$ will be applied, based on an inhomogeneous Poisson model. In addition to yielding approximations to the peak ambiguity probability, the Poisson characterization will allow a simple expression for the global variance of the time delay estimate to be derived, involving moments of the intensity function of the upcrossing count process N.

The next section is concerned with the distribution of the event $N > 0$. Section V will contain the ensuing derivation of variance, and in Section VI specific expressions for the conditional rate function (first moment of N) are derived under suitable distributional assumptions on R^g.

IV. The distribution of N

In what follows we will consider the large error interval of interest, I, to be $[t_o, t]$ for notational simplicity, e.g. $[t_o, t] = [-D_m, -\delta]$. Drop the reference to the interval I in the notation for the number of upcrossings of $-\epsilon$ by ΔR^g over the interval I, $N(I)$. Now define the conditional expectation of the k-th factorial product of N given $\Delta R^g < -\epsilon$

$$\beta_k = \mathbf{E}_c [N(N-1)...(N-k+1)] \tag{18}$$

The β_k's are commonly termed the factorial moments (conditional here) of N. Assume β_k is finite for all k and absolutely continuous with respect to intervals of time. Then it is stated in [16] that the (conditional) probability that N is greater than zero obeys the exact relation

$$P_c(N > 0) = \sum_{k=1}^{\infty} (-1)^{k-1} \frac{(\beta_k')}{k!} \tag{19}$$

where β_k' is the derivative of β_k. In particular β_1' is the conditional intensity

λ_c of N, or with $N(t)$ the notation for $N(t_o,t)$

$$\lambda_c(\tau) = \beta_1' = \frac{d\mathbf{E}_c[N(\tau)]}{d\tau} \quad , \qquad \tau \in [t_o,t] \tag{20}$$

Since $\lambda_c(\tau)\Delta$ is the expected number of upcrossings over a small interval of length Δ, the conditional intensity has the interpretation of the average number of upcrossings per unit time.

It is to be expected from Eq. (19) that λ_c plays a major role in determining the probability of upcrossings. Indeed, it was shown in the previous chapter that under certain conditions (Leadbetter's conditions given in Chapter 4 , Eqs. (6) - (8)) the (conditional) probability

$$p(t) = P_c(N(t) > 0) \tag{21}$$

satisfies the integral equation

$$p(t) = \int_{t_o}^{t} \lambda_c(\tau)(1 - p(\tau))d\tau + Q(t) \tag{22}$$

$Q(t)$ in (22) is given by

$$Q(t) = \lim_{n \to \infty} \sum_{i=0}^{2^n-1} q(t_i) \tag{23}$$

$$\tag{24}$$

$$q(t_i) = P(N(t_i,t_{i+1})>0, \ N(t_i)=0) - P(N(t_i,t_{i+1})>0)P(N(t_i)=0)$$

Here $\{t_i\}_{i=0}^{2^n}$ is an increasingly dense partition of the interval $[t_o,t]$. In Chapter 4 it was shown that the (conditional) intensity is given by the integral

$$\lambda_c(\tau) = \int_0^\infty z \ p_r(0,z)dz \tag{25}$$

where p_r is given in Appendix E as the limit of a function of the three-dimensional density of ΔR^g (This limit is the (conditional) joint density of $\Delta R^g(\tau)$ and $d\Delta R^g/d\tau$ if the derivative exists).

In Chapter 4 we went on to show that if the possibly non-stationary process ΔR^g is uniform mixing; the present becomes independent from the past rather quickly, and the upcrossings are made rare in a specific sense, then the distribution of a normalized version of the number of upcrossings converges to the Poisson law. That is, let N^* be a Poisson counting process with the identical (conditional) intensity as N (conditioned on $\Delta R^g(t_o) < -\epsilon$). The probability of getting k points in the interval of interest by this Poisson process is

$$P(N^*(t) = k) = \frac{\left[\int_{t_o}^{t} \lambda_c(\tau)d\tau\right]^k}{k!} \exp\{-\int_{t_o}^{t} \lambda_c(\tau)d\tau\} \tag{26}$$

The integral $\int_{t_o}^{t} \lambda_c(\tau)d\tau = \mathbf{E}_c[N]$ is commonly referred to as the (conditional) rate function of the point process N.

The asymptotic result mentioned above asserts that $p(t)$, Eq. (21), converges to $p^*(t)$, the probability of getting at least one point by N^* in $[t_o, t]$

$$p^*(\tau) = 1 - \exp\{-\int_{t_o}^{r} \lambda_c(\sigma)d\sigma\} \tag{27}$$

As a practical tool the above asymptotic result is unsatisfactory, in so far as the approximation of $p(t)$ by $p^*(t)$ is concerned, since it gives no indication of the agreement between the Poisson and the actual probability for finite observation time. However, Lemma 2.2 of Chapter 4 states that the difference between the Poisson and the actual upcrossing probabilities satisfies

$$| p(t) - p^*(t)| \leq Q(t) \, exp(\int_{0}^{t} h(\tau)d\tau) \tag{28}$$

Now $Q(t)$, Eq. (23), is a measure of the dependency of the present on the past of the random process. Thus if the upcrossing process $N(\sigma,\nu)$ has little dependence

on

$N(u,v)$ where $[\sigma,\nu]$ and $[u,v]$ are intervals separated by a finite distance, the error, upper bounded by Eq. (28), is arguably small.

By using the definition

$$\mathbf{E}_c [N] = \sum_{k=1}^{\infty} P_c (N = k) \tag{29}$$

we have the result that for any count process N

$$P_c (N > 0) \leq \mathbf{E}_c [N] \tag{30}$$

so that the difference $|\, p(t) - p^*(t)|$ satisfies

$$|\, p(t) - p^*(t)| \leq \mathbf{E}_c [N(t)] \tag{31}$$

Hence at the very least the approximation of $p(t)$ by $p^*(t)$ is indicative of the SNR region where the ambiguity error must be small. Indeed, in Sections VII VIII, it will be seen that the conditional rate function, and hence $p^*(t)$, undergoes a rapid increase from zero as the input SNR , relative to the observations, Chapter 2 Eq. (1), decreases beyond a certain threshold. In Section VIII it is observed that this threshold agrees with the experimentally observed threshold, for a specific broadband signal spectrum, beyond which the variance of the time delay estimate deteriorates.

However one must take into account an aspect of the Poisson model for the upcrossings which may cause overly pessimistic predictions of system performance. It is conjectured that the inaccuracy of the independent increment assumption neccessary to send $Q(t)$ to zero, in (28), results in under-estimation of the conditional variance of N, even though N^* has the identical conditional mean as N. This can be shown straightforwardly for the unconditioned mean $\mathbf{E}[N]$, for flat broadband lowpass signal spectra, by demonstrating that $var[N]$ is lower bounded by $\mathbf{E}[N]$. This is not essential for the following development and

the proof is omitted. Since the variance of the Poisson process with mean $\mathbf{E}[N]$ achieves the lower bound, the Poisson model can be viewed as that model, matching N in mean, which minimizes the variance over all other such models. Keeping in mind that $P(N = 0)$ is the quantity of interest here, the Poisson model entails a more rapid decrease in $P(N = 0)$, as $\mathbf{E}[N]$ increases, than would normally be the case. Hence $P(N > 0)$ is over-estimated by $P(N^* > 0)$.

In Section VIII the performance of a modification of the Poisson approximation, acheived through a simple normalization of λ_c, will be investigated and shown to be in better agreement with experiment for broadband signals than the Poisson approximation presented above. For broadband signals of one sided bandwidth B_o Hz , a downward correction of the intensity function is made by division of λ_c by $\sqrt{B_o}$. This has a smoothing effect on the transition of $p^*(t)$, Eq. (27), as $\mathbf{E}_c[N]$ increases from zero. The divisor was initially motivated by a numerical comparison of the Poisson approximation with Ianniello's approximation to the probability of anomaly, for large BT. However, the above modification can be regarded as assuming the following specific form for the correction term, $q(\tau)$ in (23) and (24)

$$q(\tau) = (\frac{1}{B_o} - 1)\lambda_c(\tau)(1 - p(\tau)) \qquad (32)$$

Recall, $q(t)$ characterizes the deviation of $p(t) = P_c(N(t_o,t) > 0)$ from the Poisson approximation $p^*(t)$. In general, at very low SNR and high BT, the conditional intensity becomes uniformly large over the large error region. This implies that the conditional rate function is much greater than one, which forces the true probability of upcrossing, $p(t)$, to be close to unity. The Poisson approximation, $p^*(t)$ also has this behavior. Likewise at high SNR the conditional rate is close to zero, so that, by (31), $p(t)$ and $p^*(t)$ have identical high SNR behavior. The assumed form for $q(t)$ reflects the low and high SNR agreement of the Poisson

approximation with the true probability (i.e. $q(t) \approx 0$ in these cases). The factor $(1/B_o - 1)$ attenuates the increase in the probability of ambiguity as B_o, and thus $\mathbf{E}_c [N]$, increases.

V. Expressions for Estimator Variance

In the previous Section a large error probability approximation was motivated, based only on the projection of the Poisson random variable, N^*, onto the Bernoulli random variable, X, defined as the indicator function of the event $N^* > 0$. That is $X = 1$, with probability $P_e = P(N^* > 0)$ and $X = 0$, with probability $1 - P_e$.

For the variance approximation we will exploit the full structure of N^*. Thus we specifically assume that $N(t)$, the number of upcrossings of $-\epsilon$ by ΔR^g within $[t_0, t]$, has a conditional distribution $P_c (\bullet)$, as in Eq. (12), identical to the distribution of $N^*(t)$, Eq. (14).

In what follows we will use the convention of denoting the total number of upcrossings over the large error region, $\mathbf{H}_\delta$, by N, whereas $N(I)$ and $N(t)$, for I an interval, t a point, retain their usual definitions, (16) and $N(t) = N(-D_m, t)$. Of course, $N = N(I_L) + N(I_R)$ in this notation, where I_L and I_R decompose $\mathbf{H}_\delta$ into left and right intervals, as in (7).

Define for the N peak ambiguities, or points, occurring in the large error region, $\mathbf{H}_\delta$

$$\omega_i = \left\{ location \ of \ the \ i^{th} \ point, \ i = 1, 2, \cdots, N \right\} \tag{33}$$

While there may be many ambiguities, there can only be a single ω_i which corresponds to the global maximum of R^g. Given an ambiguity has occurred, we model the global maximum of R^g, $\hat{D}$, as being equally likely to fall on any

one of the ω_i's. That is

$$P(\hat{D} \in [t, t + \Delta] \mid \omega_i \in [t, t + \Delta], \text{ some } i = 1, 2, \cdots, N) = \frac{1}{N} \qquad (34)$$

Note the above model restricts the location of the ambiguous peak of R^g to lie within Δ of one of the upcrossing points $\{\omega_i\}$. This is a good approximation for large *a priori* interval-bandwidth product, $D_m B$, when excursions of ΔR^g above $-\epsilon$ are short lived (spurious peaks).

Using the law of total probability, we can express the variance of the time delay estimate $\hat{D}$ as

$$var(\hat{D}) = var(\hat{D} \mid \hat{D} \notin \mathbf{H}_\delta) \; P(\hat{D} \notin \mathbf{H}_\delta) \qquad (35)$$
$$+ \lim_{\Delta \to 0} \sum_{t_i \in \mathbf{H}_\delta} (t_i - D_o)^2 \; P(\hat{D} \in [t_i, t_i + \Delta])$$

Here $\{t_i\}_{i=1}^{M(\Delta)}$ is an increasingly dense partition of $\mathbf{H}_\delta$ as Δ decreases to zero. The conditional variance, $var(\hat{D}|\hat{D} \notin \mathbf{H}_\delta)$, in Eq. (35) is equal to the local variance of the delay estimate derived in [2], as long as $var(\hat{D}|\hat{D} \notin \mathbf{H}_\delta) < \delta^2/3$, and is equal to $\delta^2/3$ otherwise. For simplicity denote $var_L(\hat{D})$ the conditional local variance $var(\hat{D} \mid \hat{D} \notin \mathbf{H}_\delta)$.

Using the approximate independence of the distribution of ambiguities of D_o (Appendix C), Eq. (35) is equivalent to

$$var(\hat{D}) = var_L(\hat{D}) \; P(\hat{D} \notin H\delta) \qquad (36)$$
$$+ \lim_{\Delta \to 0} \sum_{t_i \in \mathbf{H}_\delta} t_i^2 \; P(\hat{D} \in [t_i, t_i + \Delta])$$

Using (33) we have

$$P(\hat{D} \in [t, t + \Delta]) \qquad (37)$$
$$= \sum_{k=1}^{\infty} P(\hat{D} \in [t, t + \Delta], \omega_i \in [t, t + \Delta] \text{ some } i = 1, \cdots, k)$$
$$= \sum_{k=1}^{\infty} \frac{1}{k} \; P(\omega_i \in [t, t + \Delta] \text{ some } i = 1, \cdots, k)$$

Define the composite intensity function $\rho_c(\tau)$, on $\mathbf{H}_\delta = I_L \cup I_R$, where I_L and

I_R are the left and right ambiguity regions defined in Eq. (7)

$$\rho_c(\tau) = \begin{cases} \dfrac{\partial}{\partial \tau} \ E_c\left[N(D_m,\tau)\right], \ \tau \in I_L \\[2mm] \dfrac{\partial}{\partial \tau} \ E_c\left[N(D_m,\tau)\right], \ \tau \in I_R \end{cases} \tag{38}$$

In (38), $N(-D_m,\tau)$, $\tau < D_m$, is the notation for the number of upcrossings of $-\epsilon$ by ΔR^g in $[\tau, D_m]$ where $\Delta R^g(s)$ is taken as tracing out its trajectory as s decreases toward zero. Thus $E_c[N(D_m,\tau)]$, $\tau \in I_R$, is the expected number of upcrossings of this time reversed ΔR^g given $\Delta R^g(D_m) < -\epsilon$.

For the GCC trajectories of interest, the symmetry of the first and second order statistics of R^g about zero (see Appendix C) implies that $E_c[N(-D_m,\tau)] = E_c[N(D_m,\tau)]$. Thus

$$\rho_c(\tau) = \begin{cases} \lambda_c(\tau) \ , \ \tau \in I_L \\ \lambda_c(-\tau) \ , \ \tau \in I_R \end{cases} \tag{39}$$

where $\lambda_c(\tau)$ is the conditional intensity defined in Eq. (20).

In Appendix H, it is derived that (if ρ_c is continuous at $t \in \mathbf{H}_\delta$)

$$P(\omega_i \in [t, t+\Delta] \text{ some } i = 1, \cdots, k) \tag{40}$$

$$= \rho_c(t) \bullet \Delta \frac{\left[\int\limits_{\mathbf{H}_\delta} \rho_c(s)ds\right]^{k-1}}{(k-1)!} \ \exp\{-\int\limits_{\mathbf{H}_\delta} \rho_c(s)ds\}$$

Hence substituting (40) into (37) and recognizing the identities

$$P(N > 0) = \sum_{k=1}^{\infty} \frac{\left[\int\limits_{\mathbf{H}_\delta} \rho_c(s)ds\right]^{k}}{k!} \ \exp\{-\int\limits_{\mathbf{H}_\delta} \rho_c(s)ds\} \tag{41}$$

$$E\{N\} = \int\limits_{\mathbf{H}_\delta} \rho_c(s)ds \tag{42}$$

we obtain

$$var(\hat{D}) = var_L(\hat{D}) \, P(N=0) + \left[\int\limits_{\mathbf{H}_\delta} s^2 \, \frac{\rho_c(s)}{E\{N\}} \ ds\right] P(N > 0) \tag{43}$$

Recognizing $P(N > 0)$ as P_e, Eq. (9), and using (39), we have the equivalent expression for variance

$$\text{var}(\hat{D}) = \text{var}_L(\hat{D})\,[1 - P_e] + \int_{-D_m}^{-\delta} s^2\,\hat{\lambda}_c(s)\,ds \; \bullet \; P_e \tag{44}$$

where we have defined the normalized intensity function

$$\hat{\lambda}_c(\tau) = \frac{\lambda_c(\tau)}{\int_{-D_m}^{\delta} \lambda_c(s)\,ds} \tag{45}$$

Eq. (44) shows that as the probability of large error increases, one must correct the local variance by adding a proportionate amount of a potentially large quantity. This quantity can be interpreted as the second moment of the probability density of ambiguities over time, $\hat{\lambda}_c$, i.e. the mean-square deviation of the magnitude of the large error.

In Section VIII, the Poisson variance approximation (44) will be implemented using a time independent approximation to $\hat{\lambda}_c$, for large $D_m B$. In this case (44) becomes

$$\text{var}(\hat{D}) = \text{var}_L(\hat{D})\,[1 - P_e] + \frac{(D_m^2 + \delta^2)}{3}\,P_e \tag{46}$$

Now as the signal-to-noise ratio decreases to zero, one would expect the variance to correspond to the case of uniformly distributed random variable on $[-D_m, D_m]$. Indeed as this SNR condition occurs

$$P_e \;\rightarrow\; \frac{D_m - \delta}{D_m} = 1 - \frac{\delta}{D_m} \tag{47}$$

and from (46)

$$\text{var}(\hat{D}) \;\rightarrow\; \delta^2/3 \;\frac{\delta}{D_m} + \frac{(D_m^2 + \delta^2)}{3}\;\frac{(D_m - \delta)}{D_m} \tag{48}$$

With $\delta = 2\pi/B$, it is readily verified that (48) becomes

$$var(\hat{D}) \;\rightarrow\; D_m^2/3 + O(\frac{1}{D_m B}) \qquad (49)$$

which is the expected result for $D_m B \gg 1$.

VI. Explicit Form of the Average Number of Peak Ambiguities

Of interest here is the form of the conditional rate function $\mathbf{E}_c[N]$, or the time varying intensity function, $\lambda_c(\tau)$, of the number of upcrossings N, associated with the ambiguity process ΔR^g. λ_c is a function of the bivariate distributions of ΔR^g conditioned on $\Delta R^g(-D_m) < -\epsilon$. The bivariate distribution of the GCC output, R^g, is not exactly known except for particularly simple cases such as sinusoidal signals and simple cross-correlation operations [13]. Even in this simple situation the resultant expressions are cumbersome. Here we will assume that R^g is adequately characterized by Gaussian statistics. This can be justified if one has a large observation interval and implements segment averaging techniques to estimate the cross-correlation function. The details are relegated to Appendix D.

In Appendix E it is shown that if the a.s. continuous Gaussian process ΔR^g satisfies some regularity conditions, involving continuity of certain first order derivatives of its covariance function and a non-degeneracy condition, then the conditions neccessary for the validity of (21) through (25) and (28) are satisfied for the conditioned process: ΔR^g given $\Delta R^g(-D_m) < -\epsilon$. Simple moment conditions on the observation spectra are given in Appendix F which guarantee the above mentioned continuity. In Appendix E the conditional intensity function of the upcrossings of $-\epsilon$ by $\Delta R^g(\tau)$ given $\Delta R^g(-D_m) < -\epsilon$, is derived as

$$\lambda_c(\tau) = K \int_0^\infty y\; \Phi(\alpha y + \theta)\; \phi(y + \Theta) dy \;, \qquad \tau \in [-D_m, -\delta] \qquad (50)$$

K, α, θ and Θ are functions of τ given by Eqs. (E.4) through (E.10) in Proposition F.1.

Unfortunately the integral in Eq. (50) is not tabulated so that the calculation of the rate function, $\mathbf{E}_c[N]$, the integral of λ_c over $[-D_m,-\delta]$, involves a two-dimensional integration over an infinite region. Although routines do exist to perform such calculations for this type of integral [10] a simple approximation to λ_c will be presented in the next Section involving only a single integration. For purposes of comparison, it will be instructive to consider Eq. (50) under similar broadband assumptions as in [3].

For large $D_m B$, and a lowpass signal spectrum, it is shown in Appendix G that K and θ are approximately constant, and α and Θ are approximately zero, over the majority of $[-D_m, -\delta]$ (A plot of $\alpha(\tau)$ is displayed in Fig. 6.1 for a typical lowpass spectrum). As a consequence, for large $D_m B$, we lose little accuracy in our computations by using a time independent approximation to the intensity, (50), in place of the exact expression. Such an approximation is derived in Appendix G for flat lowpass signal and noise spectra of magnitudes S/B and N/B respectively, where $B/2$ is the cutoff frequency beyond which the spectra are zero.

Recall that the resolution parameter, ϵ, is related to the quantization of R^g, associated with the peak searching operation. If ϵ is not zero then it is clear, from the development in Appendix G, that in addition to involving SNR the intensity, (50), is also a decreasing function of signal and noise power. The interpretation is that resolution (quantization error) depends on the absolute magnitude of the GCC output, while the true location of the global peak only depends on the relative magnitude (i.e. it is scale independent). In theory the effect of ϵ on performance can be minimized by scaling up the GCC trajectory before quantization. In practice, however, scaling (amplification) is not a

distortion free operation and could introduce additional error. Nonetheless, we will concentrate on the case where ϵ is zero in the following discussion.

The conditional rate function of the upcrossings is the integral of the intensity, which, from (G.54), is of the form

$$\mathbf{E}_c\,[N] = \lambda_c\,(D_m - \delta) = \frac{B_o\,(D_m - \delta)}{\sqrt{3}\,\sqrt{2 + |\gamma_{12}|^2}}\;\frac{\Phi\left(\dfrac{\gamma'\,\sqrt{2B_o\,T}}{\sqrt{3 + 2|\gamma_{12}|^2}}\right)}{\Phi(\gamma'\,\sqrt{2B_o\,T})}\;\phi(\gamma'\,\sqrt{2B_o\,T}) \tag{51}$$

(Note in practice $\delta \ll D_m$ in (51)) Here $B_o = B/4\pi$ is the single sided signal bandwidth in Hz, $|\gamma_{12}|$ is the magnitude coherency

$$|\gamma_{12}| = \frac{S}{S + N} \tag{52}$$

and we have defined

$$\gamma' = \frac{|\gamma_{12}|}{\sqrt{2 + |\gamma_{12}|^2}} \tag{53}$$

The consideration of (51) gives some insight into the large error performance of the GCC. Note that, as expected, decreasing the SNR, or equivalently decreasing $|\gamma_{12}|$ and γ' towards zero, increases the intensity function, λ_c, and thus the probability of peak ambiguity, P_e. However the rate of increase in λ_c, due to decreasing SNR, is determined by the time-bandwidth product $B_o\,T$ and the *a priori*-interval-bandwidth product $B_o\,D_m$. Specifically, for fixed D_m, $B_o\,T$ determines the signal-to-noise ratio threshold beyond which increasing SNR leads to a rapid decrease in λ_c. In general, increasing $B_o\,T$ shifts this threshold to a lower SNR, expanding the range of error free operation of the GCC. On the other hand, $B_o\,D_m$ affects the magnitude of the expected number of upcrossings, and its rate of increase, as the SNR decreases beyond the threshold. Increasing

the length of the *a priori* region, $[-D_m, D_m]$, naturally makes large errors more likely as is indicated by the corresponding increase in $\mathbf{E}_c[\mathcal{N}]$ of (51) The above comments corroborate analytical results obtained in [4] and the generally acknowledged practical performance of the GCC [3].

Consider the asymptote of $\mathbf{E}_c[\mathcal{N}]$ as $|\gamma_{12}|$ goes to zero. In the limiting case , from Eq. (51)

$$\mathbf{E}_c[\mathcal{N}] \to \frac{B_o(D_m - \delta)}{\sqrt{6}} \tag{54}$$

$\mathbf{E}_c[\mathcal{N}]$ is finite if $D_m - \delta$ is non-zero so that there is a finite probability that no large error will occur even at infinitessimally small SNR. This is merely an effect of the uniform distribution of the global peaks of R^g over the *a priori* interval when no signal component is present in the observations. Note, in general, the probability of ambiguity is upper bounded by $(D_m - \delta)/D_m$, the probability of error for an estimator which ignores the observations, generating random values within $[-D_m, D_m]$.

VII. The Unconditioned Approximation to $\mathbf{E}_c[\mathcal{N}]$

Section VI dealt with the form of the conditioned rate, $\mathbf{E}_c[\mathcal{N}]$, of the upcrossings of $-\epsilon$ by ΔR^g over $[-D_m, -\delta]$ given $\Delta R^g(-D_m) < -\epsilon$. Except for large $D_m B_o$, where an explicit expression for the conditional rate function can be derived, the integral of (51), that is $\mathbf{E}_c[\mathcal{N}]$, (50), is difficult to evaluate. Here we will consider the unconditioned version of the conditional rate, $\mathbf{E}[\mathcal{N}]$, which has a simpler functional form than (50), but remains a good approximation to $\mathbf{E}_c[\mathcal{N}]$.

Recall, from Section II, conditioning of the upcrossing probability over $[-D_m, -\delta]$, on the the non-exceedance of $-\epsilon$ by $\Delta R^g(-D_m)$, was necessary to assert an equivalence between level crossings and peak ambiguities. However the difference between the subset of realizations of R^g which give peak ambiguity,

that is simply exceed $R^g(0)-\epsilon$, and those that actually upcross $R^g(0)-\epsilon$ is the subset where R^g exceeds the above level in the neighborhood of $-D_m$ and down-crosses the level at most once over $[-D_m, -\delta]$. It can be argued that the probability of this latter event is small when D_m is sufficiently large. Essentially, at high SNR, any ambiguities which may occur are predisposed to occur near the sidelobes of the cross-correlation function (where the mean value of R^g approaches the mean value of $R^g(0)$). On the other hand, at low SNR, many peak ambiguities are likely to occur, so that an upcrossing must exist within $[-D_m, -\delta]$ regardless of the value of $R^g(-D_m)$.

The unconditioned rate function, $\mathbf{E}[N]$, is derived in Reference [17] for a non-stationary Gaussian random process. The result is

$$\mathbf{E}[N] = \int_{-D_m}^{-\delta} \lambda(\tau)d\tau \tag{55}$$

where λ is the unconditioned intensity function of N

$$\lambda(\tau) = (1 - \rho_\tau^2)^{1/2}\frac{\dot{\sigma}_\tau}{\sigma_\tau}\quad \phi(\frac{\epsilon-\mu_\tau}{\sigma_\tau})\left[\phi(\varsigma_\tau) + \varsigma_\tau\Phi(\varsigma_\tau)\right] \tag{56}$$

σ_τ^2, $\dot{\sigma}_\tau^2$ and ρ_τ are: the variance of $\Delta R^g(\tau)$, the second derivative of the variance, and the first derivative of the variance evaluated at time τ. μ_τ and η_τ are the mean of $\Delta R^g(\tau)$ and its derivative. ς_τ is the ratio

$$\varsigma_\tau = \frac{\eta_\tau - \rho_\tau\dfrac{\dot{\sigma}_\tau}{\sigma_\tau}(\epsilon + \mu_\tau)}{(1 - \rho_\tau^2)^{1/2}\dot{\sigma}_\tau} \tag{57}$$

Again assume $\epsilon = 0$. An asymptotic analysis of identical form as that applied to λ_c and $\mathbf{E}_c[N]$ in Section VI to obtain Eq. (51), gives the following time independent approximation to the unconditioned rate

$$\mathbf{E}[N] = \lambda(D_m - \delta) = \frac{B_o(D_m - \delta)}{\sqrt{3}\sqrt{2+|\gamma_{12}|^2}}\sqrt{2\pi}\,\phi(\gamma'\sqrt{2B_o T}) \tag{58}$$

where $|\gamma_{12}|$, γ' and B_o are defined as in Eq. (51). The unconditional approximation and the true rate function, Eq. (58) and Eq. (51), differ only in the absence, in (51), of the ratio $\Phi(ax)/\Phi(x)$, with $a = \sqrt{2+|\gamma_{12}|^2}$ and $x = \gamma' \sqrt{2B_o T}$. This ratio is essentially unity for large and small SNR ($\gamma' \sqrt{2B_o T}$), which is consistent with the argument at the beginning of this section concerning the accuracy of the approximation of $\mathbf{E}_c [N]$ by $\mathbf{E}[N]$. For comparison, plots of the rate function, $\mathbf{E}[N]$, and the conditional rate function, $\mathbf{E}_c [N]$, are displayed in Fig. 8.1 for $B_o = 200Hz$ $B_o = 100Hz.$ and $B_o = 50Hz$ for $D_m = 0.125sec.$, $T = 8secs$. The unconditional approximation is uniformly larger than the quantity it is attempting to estimate, i.e. (51). This indicates that, at least for large $D_m B$, at worst, the Poisson probability approximation, using (56) in place of (50), will over-estimate the nominal approximation, using (50), below the SNR threshold.

In the next section the unconditional rate of (55) is used in place of the conditional rate of (50) to numerically evaluate the ambiguity probability and error variance approximations, obtained by applying the Poisson approximation to the conditional probability of upcrossings, discussed in Sections IV and V. For reference we display the specific approximation to the probability of ambiguity, P_e of (10) and (12), for broadband and narrowband signal spectra respectively

$$P_e = P(A_{I_L})(2 - P(A_{I_L}))$$ (59)

and

$$P_e = P(A_{I_L})$$ (60)

$P(A_{I_L})$ is the probability that an ambiguity occurs over the interval $I_L = [-D_m ,\delta]$. This is a function of the conditional probability of upcrossings, $P_c (A_{I_L})$, (see Eq. (17)) which we replace by the Poisson approximation, $p^*(t)$ of (27), with λ replacing λ_c, (Note we use in (27): $[t_o ,t] = [-D_m ,-\delta]$):

$$P(A_{I_L}) = (1 - \exp\{-\mathbf{E}[N]\})P(\Delta R^{\,g}(-D_m) < 0) + 1 - P(\Delta R^{\,g}(-D_m) < 0) \quad (61)$$

where

$$P(\Delta R^{\,g}(-D_m) < 0) = \Phi\left(-\frac{\mu(-D_m)}{\sigma(-D_m)}\right) \quad (62)$$

Eqs. (59)-(61) will be referred to in the next section as the "Poisson probability approximations" and the variance obtained by substitution of P_e into Eq. (44), will be the "Poisson variance approximation".

VIII. Numerical Study

Here plots of the unconditioned intensity, Eq. (56), the rate function, Eq. (55), the Poisson probability approximations, Eqs. (59)-(61), and the Poisson variance approximation, Eq. (44), are investigated for broadband and narrowband spectra. First we consider the case of broadband lowpass signal spectra with large BT. For this case we use, as a benchmark for comparison, the specific signal and observation parameters as were used in [14]. For the example in [14] we plot the intensity surface as a function of time and SNR to gain some insight into the distribution of large errors over the *a priori* region of delay. Next curves for the rate function of ambiguity and the resultant Poisson approximations are displayed, and the probability of large error is compared against Ianniello's approximation, (2). Also displayed is the corrected Poisson approximation, discussed in Section IV, designed to compensate for the Poisson models tendency to over-estimate the peak ambiguity probability. For comparison, the results of a simulation of estimator performance, undertaken in [14], is plotted alongside of the Poisson variance approximation. Finally narrowband signal spectra are considered. The intensity surface of the ambiguities is generated and compared to that generated for lowpass signal spectra. Then, based on plots of the rate function, the influences on performance of center frequency, bandwidth, observation

time and *a priori* interval are studied. The Poisson probability and variance approximations are investigated and the indicated SNR threshold values are compared to the predictions of [18], which were obtained by the Ziv-Zakai bound.

Fig. 8.2 is a plot of the intensity surface as a function of input SNR and time for a flat lowpass signal spectrum of single sided bandwidth, $B_o = 100 Hz$. The observation time was set to $T=8secs$. The time window displayed in the figures extends across the first few sidelobes of the auto-correlation function of the signal, starting at $\delta = 1/2B_o$ (at the north-east corner of Fig. 8.2 for a SNR of 0dB). In the figure, the global maximum of the signal auto-correlation function lies beyond the rightmost point on the t axis. Note the abrupt increase of the intensity near the SNR threshold indicated by SNR_t on Fig. 8.2. At low SNR the intensity surface saturates (average height is about 4.2 ambiguities per unit time) with only a small ripple over time. For high SNR the intensity is essentially zero uniformly over the time window. These correspond to the cases when the uniform distribution of peak ambiguities over the *a priori* region is an accurate approximation. That is, when Ianniello's approximation, (2), and the asymptotic forms of the intensity, (51) and (58), are strictly accurate. However in the "ambiguity dominated" region of SNR (approximately -8dB to -20dB), ambiguity occurs more frequently in the neighborhoods of the local maxima of the signal auto-correlation.

In Fig. 8.3 the normalized difference between the rate function computed from $\lambda(\tau)$, (56), and the time independent approximation to $\lambda(\tau)$, λ of (58),

$$difference = \frac{\lambda(D_m - \delta) - \int_{-D_m}^{-\delta} \lambda(\tau)d\tau}{\int_{-D_m}^{-\delta} \lambda(\tau)d\tau} \qquad (63)$$

is displayed for lowpass signals with $B_o = 100Hz$, $T = 8secs.$ and $D_m = 0.125$

and 0.02 seconds respectively. Note the above choices for D_m represent an a *priori*-interval-bandwidth product of 12.5 and 2 respectively. It is evident from Fig. 8.3 that only a small error is incurred for the larger of the $D_m B_o$ products, the error increasing as $D_m B_o$ becomes small (in which case the GCC trajectory is increasingly dominated by the signal auto-correlation sidelobe activity over the a *priori* region). The above observations provide evidence for the fact that under the uniform assumption the accuracy of probability statements, concerning the number of ambiguities over $[-D_m, -\delta]$, is determined by the a *priori*-interval-bandwidth product, $D_m B_o$

Next we investigate the Poisson motivated approximations to the probability of peak ambiguity and estimator variance, discussed in Sections VI, Eq. (59) and (61), and V, Eq. (44), , for simple flat lowpass signal spectra as above. The approximation to P_e of (59) is plotted in Figs. 8.4, 8.5 and 8.6 against Ianniello's formula, (2), as a function of input SNR for $B_o = 50Hz$, $B_o = 100Hz$, and $B_o = 600Hz$ respectively, and $T = 8secs.$, $D_m = 0.125secs$. Also plotted is the corrected version of the Poisson probability approximation discussed in Section IV. The Poisson probability approximation over-estimates the ambiguity probability within the regions $-6dB$ to $-30dB$, $-8dB$ to $-30dB$ and $-12dB$ to $-20dB$ for $B_o = 50Hz, B_o = 100Hz$ and $B_o = 600Hz$ respectively. The error in the Poisson approximation is most severe near the knee of Ianniello's curve and increases as B_o decreases. This is in agreement with the comments on the behavior of the Poisson approximation at the end of Section IV. Note, however, that the Poisson model gives good approximations to the locations of the SNR thresholds, which are within 2dB of those indicated by Ianniello's approximation. The corrected Poisson, implemented using $\mathbf{E}_c[N]/\sqrt{B_o}$ in place of $\mathbf{E}_c[N]$, in Eqs. (27), (14) and (10), gives better agreement with Ianniello's probability approximation as is indicated in Figs. 8.4 through 8.6.

In Fig. 8.7 the variance approximation of Section V is compared to the CRLB (local variance of [2]), and to the simulated variances calculated in [14], using the Poisson and corrected Poisson approximations, for $B_o = 100Hz$, $T = 8secs.$, and $D_m = 0.125secs$. The Poisson approximations agree with the CRLB above a SNR of about -6dB. However, as the SNR falls below this threshold, the Poisson approximations indicate a much higher variance than that indicated by the CRLB. This is the ambiguity dominated region, where the CRLB fails to govern the performance of the GCC. Note that the general behavior of the Poisson approximations agree well with the simulated performance. However, the nominal Poisson approximation gives variance predictions which are somewhat conservative over the threshold region, $-6dB$ to $-15dB$, while the corrected Poisson model for the variance is closer, on the average, to experiment.

Figs. 8.8 through 8.18 relate to the bandpass signal spectra. General comments concerning the Poisson approximations are as follows. For the bandpass case, the auto-correlation of the signal has a very narrow peak (lobe) at the origin. Therefore, at high SNR the estimate of the location of this peak is highly accurate (for high BT), much more so than for the lowpass case, as is predicted by the CRLB. However while the local variance indicates a basically log linear decrease in performance as the SNR decreases, the Poisson approximation reveals the SNR thresholds discussed by Weiss and Weinstein in [18] which are ignored by the local variance.

In Fig. 8.8 the intensity surface, as a function of time and SNR, is displayed for a bandpass signal at center frequency $f_o = 500Hz$, with one sided bandwidth $B_o = 100Hz$, and $T = 8.0secs$. Here the time interval extends from the first zero crossing of the auto-correlation function of the signal, $\delta = 1/4f_o$, to approximately the fifth sidelobe away from the origin. Fig. 8.8 is oriented identically to

Fig. 8.2 for the lowpass case , i.e. the global maximum of the autocorrelation is beyond the rightmost point on the t axis. Contrasting Fig. 8.8 with Fig. 8.2 it is evident that the variation in the intensity surface over time is much more severe in the bandpass case, even at low SNR. In fact the average number of ambiguities is orders of magnitude greater near the first sidelobe than over the rest of the time axis. A distinctive feature of Fig. 8.8 is the SNR difference between the point, SNR_1 where a rapid rise in the intensity of ambiguity first begins, i.e. in the region of the first sidelobe, and the point, SNR_2 where a uniform increase of the ambiguity, over time, is in evidence. This implies the existence of at least two separate SNR thresholds in the bandpass case.

In [18] four distinct regions of performance were discovered based on a study of the Ziv-Zakai lower bound. These regions are delineated by two SNR thresholds, SNR_1 and SNR_2, and a SNR point, SNR_3 , beyond which only *a priori* information is useful. The Ziv-Zakai bound can be loosely interpreted as governing the performance of the optimal estimator of delay (The bound, however, is generally unattainable). In light of the present results, displayed in Fig. 8.8, a physical explanation of the thresholds can be proposed. It may be helpful to refer to Fig. 2.2a and Fig. 2.2b in interpreting the following comments. For SNR larger than SNR_1, the only significant source of errors comes from small variations in the maximum of the narrow peak, which occurs at the true delay. As the SNR approaches SNR_1, however, a rapid increase in the error occurs, due to the proximity of the maxima of the closely spaced high frequency sidelobes of the signal auto-correlation. Due to the rapid attenuation of the high frequency component by the auto-correlation envelope, Fig. 2.3b, an initial saturation of the errors occurs within the central lobe of the envelope as the SNR approaches SNR_2. The occurrence of additional large errors, over the outlying remainder of the *a priori* interval, is precluded until a sufficiently low second threshold is

attained, SNR_2. Beyond SNR_2, the outlying ambiguity becomes as significant as that falling within the central lobe of the envelope, and the error begins a second period of rapid increase. This increase continues until total saturation of the *a priori* interval is achieved at SNR_3. Beyond SNR_3, a limit on the number of ambiguities is imposed by the maximum number of times a waveform, of finite bandwidth B_o, can undergo zero crossings within the *a priori* region. While SNR_3 is not really a threshold in the sense of SNR_2 and SNR_1, in the following we will refer to all three as SNR thresholds, for ease in presentation.

The remaining figures indicate the influence of the parameters B_o, f_o, T and D_m on the three quantities: the rate function of peak ambiguity, and the Poisson probability and variance approximations. In Fig. 8.9 a table of values of the thresholds, SNR_3, SNR_2 and SNR_1, predicted by the theoretical analysis of the Ziv-Zakai lower bound [18 Eqs. (62), (72) and (73)] is presented for comparison with the thresholds indicated by the Poisson approximation in Figs. 8.12, 8.15, and 8.18.

In Fig. 8.10 through Fig. 8.12 the above three quantities are plotted as a function of SNR, for fixed parameters: $f_o = 2000Hz$, $T = 8.0secs.$ and $D_m = 0.125secs.$ and varying bandwidth: $B_o = 100Hz$ and $500Hz$. In addition, in Fig. 8.12 the CRLB is plotted for comparison. We note the following features. An increase in B_o entails a decrease in the rate function of ambiguity and a decrease in the initial SNR threshold SNR_1. At low SNR, the steady state magnitude of the rate function does not significantly vary over the two bandwidths. For $B_o = 500Hz$ (BT=4000), as we increase the SNR the variance approximation joins the CRLB at $SNR \approx -6dB$. However, for $B_o = 100Hz$ (BT=1600), it appears to hit an asymptote of constant non-zero error. This effect was also reported in [19] by applying a small error approximation to the variance for bandpass signals. Specifically, for low BT, it was found that even at infinite

SNR some ambiguity will occur at the secondary sidelobes of the signal auto-correlation function. Thus, for low BT, the GCC is not an optimal estimator for bandpass signals. This irreducible error was reported to be removable by implementing a simple "symmetry checker" at high SNR in [19].

In Figs. 8.13 through 8.15, we fix the parameters: $f_o = 2000Hz$, $B_o = 100Hz$ and $D_m = 0.01$, while T is varied from 0.08sec. to 8.0sec. As in the previous case, the average number of ambiguities remains essentially unchanged at low SNR. The thresholds on the variance, and those on the probability approximations, move to lower SNR as T is increased. This is physically justified since the acquisition of more data (observation time) should yield estimates with better statistical properties. As above, an irreducible error at high SNR is indicated for low BT (T=0.08sec.). However, it is much more severe here where $BT = 8$.

Finally, Figs. 8.16 through 8.18 show the results of changing the center frequency, $f_o = 2000Hz$ and $f_o = 500Hz$, while we fix the parameters: $B_o = 100Hz, D_m = 0.125, T = 8.0sec.$. The rate parameter differs markedly in its steady state value at low SNR between the two cases. This is because the high frequency dynamics of the auto-correlation function are the limiting factor on the average number of ambiguities over a fixed interval. The general movement of the thresholds is toward larger SNRs as we decrease the center frequency, which is consistent with the Ziv-Zakai bound's behavior reported in [18]. Thus, while the local variance predicts more favorable performance at high frequencies, the Poisson variance approximation brings out the opposite fact, as the SNR falls below threshold.

From Figs. 8.12 and 8.14 it is evident, from the curves with parameters : $D_m = 0.125sec.$, $T = 8.0sec.$, $B_o = 200Hz, f_o = 2000Hz$ (Fig. 8.12) and $D_m = 0.01sec.$, $T = 8.0sec.$, $B_o = 200Hz, f_o = 2000Hz$ (Fig. 8.14), that the

SNR thresholds remain unchanged. However, the point where *a priori* information dominates the error, SNR_3, naturally differs in the above two cases. Hence, as predicted in [18], SNR_1 and SNR_2 are independent of the *a priori* interval of time delay.

Comparison of the table of threshold values, in Fig. 8.9, with the observed thresholds, using the Poisson approximations, makes a few comments neccessary. In general, for low BT, the two lower thresholds, SNR_2 and SNR_3, of the Poisson variance approximation occur very near the thresholds predicted by the Ziv-Zakai bound. The initial threshold, SNR_1, of the approximation occurs at much higher SNR than the theoretical prediction. Again, this latter behavior corroborates the conclusions of [19], where the experimental performance of the bandpass GCC is unable to achieve the CRLB near threshold. For high BT the opposite situation is the case. The Poisson approximation agrees with the Ziv-Zakai theory up to the initial SNR threshold, SNR_1, but is overly pessimistic in as far as SNR_2 and SNR_3 are concerned. This may indicate that while the high BT GCC is optimal for bandpass spectra, the low BT GCC is not. On the other hand, this behavior may be due to the inaccuracy of the Poisson approximation below threshold, as was discussed in Section IV. These conjectures can only be answered by extensive simulation studies.

IX. Discussion

In the previous sections a general framework for peak ambiguity probability in time delay estimation was developed in terms of level crossing probabilities. The representation of level crossing probabilities, using only the intensity function of the level crossings over the *a priori* interval and a correction term, led naturally to a Poisson approximation to the probability of large error and the global variance of the estimate. An analytic expression for the expected number of ambiguities (upcrossings of the ambiguity process) was derived under a

Gaussian model for the cross-correlation. This allowed a magnitude resolution parameter to be associated with the peak detection algorithm which is then applied to the cross-correlation trajectory. It was established that, under imperfect magnitude resolution, the average number of ambiguities is a function of the signal and noise powers in addition to SNR. This is indicative of the difficulty in detecting small differences of magnitude in quantized versions of weak or low magnitude waveforms. The result suggests that (noiseless) amplification of the correlation trajectory can control this type of resolution error.

For lowpass signals, a simple expression governing the average number of ambiguities over a large *a priori* region of delay was derived. The resulting expression behaves in a manner consistent with previous studies of the performance trade-offs that exist between SNR and time- bandwidth product. Not so explicitly identified elsewhere is the role of the *a priori*-interval-bandwidth product in determining performance. Indeed while the time-bandwidth product determines the location of the SNR threshold, beyond which large errors predominate estimator error, the *a priori*-interval-bandwidth product controls the rapidity with which the performance degenerates beyond the threshold.

The exact intensity function of the peak ambiguities was then investigated for both lowpass and bandpass signals. The analysis revealed that, while for high SNR the intensity is uniformly small, for intermediate SNR (i.e. within the ambiguity dominated region) the distribution of the peak ambiguities over time is far from uniform, contrary to what was assumed in [3]. The most severe variation of the intensity over time occurs in the neighborhood of the local maxima on the sidelobes of the signal auto-correlation function. The effect is most prominent for the bandpass signal, but still significant for the lowpass signal over a wide range of SNR. The degree to which the inaccuracy of the uniform assumption affects the time independent approximation of the average number of ambiguities over

the *a priori* interval, for a lowpass signal, depends on the length of the *a priori* interval relative to the correlation time of the signal, i.e. *a priori*-interval-bandwidth product. In general, a large *a priori* region is essential for the sidelobe activity of the cross-correlation to be neglected.

The bandpass intensity surface revealed the presence of four SNR regions where the behavior of the ambiguity is distinctly different. Over the first region, which extends from SNR_1 to infinity, the intensity is uniformly small over the *a priori* interval. The other three regions can be called performance breakdown regions. In the first breakdown region, demarcated by two SNR thresholds, SNR_1 and SNR_2, the error is dominated by ambiguities which occur within the central lobe of the envelope of the bandpass signal auto-correlation function. This can be interpreted as a second small error region of SNR, that complements the small error region within the central high frequency lobe (The linear region for the local variance analysis of Schultheiss [2]). As the SNR decreases beyond the first breakdown region, one enters a region where errors become significant outside of the central lobe of the envelope. Finally, as the SNR decreases further, the saturation region is attained where errors occur uniformly over the entire *a priori* interval. The existence of multiple SNR thresholds supports the work of Weiss and Weinstein [18] on the behavior of the Ziv-Zakai lower bound on estimator variance.

While a series expansion of the exact ambiguity (upcrossing) probability is known to exist, it requires knowledge of all of the higher moments of the number of ambiguities (upcrossings). Here first moment characterizations of the peak ambiguity probability, the Poisson and corrected Poisson models, in addition to the associated expressions for global variance, were investigated and compared to Ianniello's probability of anomaly for broadband lowpass signals and large *a priori*-interval-bandwidth products. The Poisson approximation over-estimates

the probability of ambiguity for lowpass signal spectra below threshold. However it well characterizes the location of the SNR threshold. The corrected version gives more accurate results for these broadband lowpass signals, but its accuracy is unknown under narrowband signal conditions.

For the low BT, bandpass case and high SNR, the Poisson variance approximation indicates the suboptimal performance of the GCC, as was reported in [19]. Likewise for high BT, the approximation attests to the optimality of the GCC for high SNR. The general behavior of the approximation, as the various parameters (observation time, center frequency, bandwidth and *a priori* interval) are varied, is consistent with the behavior of the Ziv-Zakai lower bound reported in [18]. In future studies the applicability of the Poisson approximations should be investigated for narrowband signals by means of comparisons between experiment and theory.

The corrected Poisson approximation can be interpreted as resulting from a rough model for the remainder term, $q(\tau)$, in the representation of the probability of ambiguity (upcrossing) (See Eqs. (22), (23) and (24)). A search for tractable tight bounds on this remainder may yield better first moment approximations to the probability of large errors than the Poisson examples considered here. Alternately, inclusion of higher order moments of the ambiguities, such as variance, may be a fruitful approach to improve upon these simple first moment methods.

Guide to Appendices

Chapter 5

Appendices A,B and C elaborate on the statements made in Section III of this chapter. In Appendix A the positive correlation between left and right ambiguity intervals in the correlation domain for narrowband signals is clarified for a simple choice of signal spectrum. In Appendix B the independence of the large error probability on the actual value of the delay parameter is proven for large *a priori* interval of delay. Appendix C deals with the questions of the shift invariance of the first and second order statistics of the GCC output relative to the true delay, and the uncorrelatedness of the GCC trajectory between the left and right ambiguity intervals for lowpass signal spectra.

Appendices D,E,F and G treat the approximation problems in Section VI. Appendix D contains a justification of the Gaussian assumption on the GCC statistics. Appendix E contains the derivation of the rate function of the boundary conditioned upcrossings. In Appendix F, sufficient conditions, in terms of spectral moments of the observation spectra, are given such that the derivative continuity, required in the derivation of Appendix E, is satisfied. In Appendix G, the final form of the uniform approximation for lowpass signals is derived.

Appendix I contains a derivation of an identity which underlies the inhomogeneous Poisson variance approximation of Section V.

APPENDIX A

Here it is shown that for x_1, x_2 observed over an interval $[0, T]$ as in Eq. (1) of Chapter 2, $D_0 = 0$ and flat bandlimited signal spectra there is a positive correlation between $R^g(\tau)$ and $R^g(-\tau)$ in the neighborhoods of the maxima of the sidelobes of $R_{ss}(\tau)$.

From Chapter 3 Appendix C the covariance between $R^g(\tau)$ and $R^g(-\tau)$ is

$$\sigma_{\tau,\,-\tau} = \sigma_{\tau,\,\tau-2\tau} = \frac{1}{2\pi\,T} \int\limits_{-\infty}^{\infty} \left[G_{11}(\omega)\,G_{22}(\omega) + |G_{12}(\omega)|^2\, e^{\,j2\tau\omega} \right]\, e^{-j2\tau\omega}\,d\omega \quad (\text{A.1})$$

where it has been assumed that the observation time, **T**, is much greater than the decorrelation time, T_c, of the signal, no segment averaging has been used, and $W(\omega) = 1$ in the GCC. We will consider the lowpass and bandpass cases separately.

For the lowpass case, assume the signal and noise spectra are flat out to $\pm B/2$, $B < \infty$, with level S/B and N/B respectively. Then Eq. (A.1) becomes after some manipulation

$$\sigma_{\tau,\,-\tau} = \frac{1}{2\pi\,BT}\ [(S+N)^2\ sinc(B\,\tau) + S^2] \quad (\text{A.2})$$

where

$$sinc(x) = \frac{sin(x)}{x} \quad (\text{A.3})$$

It is easily verified that the auto-correlation function R_{ss} has the form

$$R_{ss}(\tau) = \frac{S}{2\pi}\ sinc(B/2\ \tau) \quad (\text{A.4})$$

First note that if the signal-to-noise ratio is sufficiently high, then from Eq. (A.2) all points τ equidistant from zero are positively correlated. The condition for this is

$$S/N > \frac{\sqrt{2/3\pi}}{1 - \sqrt{2/3\pi}}\ \approx 0.85 \quad (\text{A.5})$$

For arbitrary signal-to-noise ratio note that at the positive peaks of $R_{ss}(\tau)$, $\tau = \dfrac{(4n-1)\,\pi}{2B}$, $n = 1, 2, \cdots$ [see Eq. (A.4)], we have from Eq. (A.2) $\sigma_{\tau,\,-\tau} = S^2/2\pi B$. Therefore, if B is small, a substantial positive correlation exists in the regions of the ambiguity prone sidelobes of $R_{ss}(\tau)$.

Now let the signal and noises be flat bandpass centered about w_s with double sided bandwidth $B < \infty$ and level $S/(2B)$, $N/(2B)$ respectively. Then from Eq. (A.1)

$$\sigma_{\tau, -\tau} = \frac{1}{2\pi BT} \ [cos(2w_s \tau) \ sinc(B\tau)(S + N)^2 + S^2] \qquad (A.6)$$

and

$$R_{ss}(\tau) = \frac{S}{2\pi} \ cos(w_s \tau) \ sinc(B/2\tau) \qquad (A.7)$$

From Eqs. (A.7) and (A.6) the identical result holds: for $\tau = \dfrac{(4n - 1)\pi}{4w_s}$ or $\tau = \dfrac{(4m - 1)\pi}{2B}$ $\quad m, n = 1, 2, \cdots \quad \sigma_{\tau, -\tau} = S^2/2\pi BT$, which is greater than zero.

APPENDIX B

Proposition

Let T_c be the dependence time of R^g and assume $R^g(\tau)$ is stationary outside of $\tau = [D_0 - T_c, D_0 + T_c]$. Then if $[-D_m, D_m]$ is the *a priori* region for D_0, and $D_m > 2T_c$, the large error probability, $P_e = P_e(D_0)$, is independent of D_0 as D_0 ranges over $(-D_m + 2T_c, D_m - 2T_c)$ under the Poisson approximation in Section IV.

Proof

Fix D_0, $|D_0| \leq D_m - T_c$ and divide $[-D_m, D_m]$ into $N = \left[\dfrac{D_m}{T_c}\right]$ intervals I_i, such that the right endpoint of one of these, I_l say, is at D_0, and $I_2, \cdots, I_{n-1}$ are of length T_c. Note by assumption $3 \leq l \leq N-3$. Hence except for intervals I_l, I_{l+1} R^g is stationary over $[-D_m, D_m]$. With $P(A_j)$ the probability that a large error occurs in interval I_j write

$$P_e(D_0) = P(\bigcup_{i=1}^{N} A_i) = 1 - P(\overline{\bigcup_{i=1}^{N} A_i}) \qquad (B.1)$$

Now use the product law for probabilities to express (B.1) as

$$P_e(D_0) = 1 - \prod_{k=2}^{N} P(\bar{A}_k \mid \bigcap_{i=1}^{k-1} \bar{A}_i) \bullet P(\bar{A}_1) \tag{B.2}$$

From the independence of A_k, A_j $|k - j| > 1$

$$P_e(D_0) = 1 - \prod_{k=2}^{N} \frac{P(\bar{A}_k \cap \bar{A}_{k-1})}{P(\bar{A}_{k-1})} \bullet P(\bar{A}_1) \tag{B.3}$$

$$= 1 - \left[\frac{1}{P(\bar{A}_{N-1})} \prod_{k=3}^{N-1} \frac{P(\bar{A}_k \cap \bar{A}_{k-1})}{P(\bar{A}_{k-1})} \right] \bullet P(\bar{A}_2 \cap \bar{A}_1) \bullet P(\bar{A}_N \cap \bar{A}_{N-1})$$

For the GCC implementations of interest, the statistics of $R^g(\tau)$ undergo only time translation as D_0 varies (see Appendix D). Hence, if $|D_0| < |D_m - 2T_c|$, $[\bullet]$ in Eq. (B.3) is independent of D_0

$$\frac{1}{P(\bar{A}_{N-1})} \prod_{k=3}^{N-1} \frac{P(\bar{A}_k \cap \bar{A}_{k-1})}{P(\bar{A}_{k-1})} \tag{B.4}$$

$$= \left[\frac{1}{P(\bar{A}_{N-1})} \prod_{\substack{k=3 \\ k \neq l,\, l+1,\, l+2}}^{N-1} \frac{P(\bar{A}_k \cap \bar{A}_{k-1})}{P(\bar{A}_{k-1})} \right] \bullet \left[\prod_{k=l}^{l+2} \frac{P(\bar{A}_k \cap \bar{A}_{k-1})}{P(\bar{A}_{k-1})} \right]$$

The first product to the right of Eq. (B.4) is a constant independent of l by the stationarity of R^g over the relevant regions. The same holds for the second product by the shift invariance of the statistics of R^g mentioned above. Hence, for any two $D_0 \in (-D_m + 2T_c, D_m - 2T_c)$, D' and D'' say, by Eq. (B.3)

$$|P_e(D') - P_e(D'')|$$

$$\tag{B.5}$$

$$= K |P(\overline{A_1' \cup A_2'}) P(\overline{A_N' \cup A_{N-1}'}) - P(\overline{A_1'' \cup A_2''}) P(\overline{A_1'' \cup A_{N-1}''})|$$

where A_j', A_j'' ; $j = 1, N$, denote events over the two extreme intervals in the partition construction for D' and D'' respectively, and K is given by Eq. (B.4).

Note if I_j' and I_j'' are the intervals over which A_j' and A_j'' are respectively defined then

$$\|I_1' \cup I_2'\| + \|I_{N-1}' \cup I_N'\| = \|I_1'' \cup I_2''\| + \|I_{N-1}'' \cup I_N''\| = L \qquad \text{(B.6)}$$

where L a common length.

For I an interval of length $\|I\|$, disjoint from $[D_o - T_c, D_o + T_c]$, the stationarity of R^g over I implies the Poisson large error approximation of (26) is only dependent on the length of I. Furthermore, substitution of the Poisson approximation to P_e, P^* of Eq. (27), into Eq. (17) gives

$$P(N(I) > 0) = \left[1 - exp(-\lambda_c \|I\|)\right] F(-\epsilon) + (1 - F(-\epsilon)) \qquad \text{(B.7)}$$

or

$$P(N(I) = 0) = exp(-\lambda_c \|I\|) F(-\epsilon) \qquad \text{(B.8)}$$

Here we have used the notation $F(-\epsilon) = F(-\epsilon, \partial^o I)$, the probability that $R^g(\partial^o)$ is less than $-\epsilon$, and $\lambda_c = \lambda_c(\tau)$, the conditional intensity; respectively independent of the endpoint $\partial^o I$ and τ by stationarity. Use of Eq. (B.8) in Eq. (B.5) for $P(A_1' \cup A_2') = P(N(I_1' \cup I_2') > 0)$, and analogously for the other three quantities, gives the desired result

$$|P_e(D') - P_e(D'')| = 0 \qquad \text{(B.9)}$$

when relation (B.6) is taken into account.

APPENDIX C

Here two statements will be clarified. First, $R^g(\tau)$ is approximately independent between the left and right large error intervals $\tau \in I_L$ and $\tau \in I_R$, given in Eq. (3.7), for lowpass broadband signals when δ in Eq. (3.7) is set to $2\pi/B$, B the two-sided bandwidth of the signal. Second, we show the shift invariance of the first and second order statistics of R^g relative to the true delay

D_0. Throughout we assume R^g to be well characterized by a Gaussian model (see Appendix E).

$R^g(\eta)$ has mean

$$E\{R^g(\eta)\} = \frac{1}{2\pi} \int_{-\infty}^{\infty} G_{12}(\omega)\, W(\omega)e^{j\omega\eta}d\omega \qquad (C.1)$$

and from Chapter 3 Appendix C $R^g(\eta)$, $R^g(\eta + \tau)$ have covariance

$$\sigma(\eta, \eta + \tau) = \frac{1}{2\pi k} \int_{-\infty}^{\infty} [G_{11}(\omega)G_{22}(\omega) + G_{12}^2(\omega)e^{j2\omega\eta}]\, W(\omega)e^{j\omega\tau} \qquad (C.2)$$

where k is proportional to T; the sensor observation time. Using the fact

$$G_{12}(\omega) = G_{\bullet\bullet}(\omega)e^{-j\omega D_0} \qquad (C.3)$$

(C.1) and (C.2) become

$$E\{R^g(\eta)\} = \frac{1}{2\pi} \int_{-\infty}^{\infty} G_{\bullet\bullet}(\omega)\, W(\omega)e^{j\omega(\eta - D_0)}d\omega \qquad (C.4)$$

$$\sigma(\eta, \eta + \tau) = \frac{1}{2\pi T} \int_{-\infty}^{\infty} \left[G_{11}(\omega)G_{22}(\omega) + G_{\bullet\bullet}(\omega)e^{j2(\eta - D_0)\omega} \right] e^{j\omega\tau}d\omega \qquad (C.5)$$

Clearly (C.4) and (C.5) depend on η and D_0 only through the difference $\eta + D_0$. Since the two first moments of a Gaussian process completely characterize its statistics the shift invariance of the statistics of R^g to D_0 is established.

Let $D_0 = 0$ and assume $s(t)$, $n_1(t)$ and $n_2(t)$ are lowpass with flat spectra of level S/B and N/B out to frequency $\pm B/2$. Then computation of the integral in Eq. (C.1) for $W(\omega) = 1$ in the passband of $s(t)$, $W(\omega) = 0$ in the stopband, gives

$$\sigma(\eta, \eta + \tau) = \frac{1}{2\pi BT} [(S + N)^2\, sinc(B/2\tau) + S^2\, sinc(\frac{B}{2}\ (2\eta + \tau))] \qquad (C.6)$$

The length of time which separates I_L and I_R is $4\pi/B$ when δ in Eq. (3.7) is chosen to be $2\pi/B$. Hence, it suffices to note that $sinc(B/2\tau) \approx 0$ for $|\eta| > \frac{2\pi}{B}$ and B large to see that R^g must be uncorrelated, and hence independent, from

I_L to I_R.

APPENDIX D

Here we show that for sufficient observation time **T** the GCC trajectory may be described to a good approximation by Gaussian statistics, if it has a finite third absolute moment (Eq. E.11).

Assume that the observation $x_i(t)$, $i = 1, 2$ becomes independent over intervals separated by T_c or more. We divide the records of length T into m subrecords each of length T_m where $T_m >> T_c$ and $T_m >> D_m$. Then form the raw cross-correlation estimates $\hat{R}_k(\tau)$, $k=1, \cdots, m$, based only on the kth subrecord.

$$\hat{R}_k(\tau) = \int_{-\infty}^{\infty} x_1^k(\sigma)\, x_2^k(\sigma + \tau) d\tau \qquad \tau \in [-T_m, T_m] \tag{D.1}$$

where

$$x_i^k(t) = \begin{cases} x_i(t) & (k-1)T_m < t \le k\, T_m \\ 0 & \text{otherwise} \end{cases} \tag{D.2}$$

The cross-correlation estimate is then obtained by averaging over the raw estimates.

$$\hat{R}_{12}(\tau) = \frac{1}{m} \sum_{k=1}^{m} \hat{R}_k(\tau) \qquad \tau \in [-T_m, T_m] \tag{D.3}$$

For T_m sufficiently large $\hat{R}_{12}(\tau)$ is represented as the sum of m independent identically distributed random quantities with the following first and second order statistics (Chapter 3 Appendix C)

$$E[\hat{R}_k(\tau)] = R_{ss}(\tau - D) \tag{D.4}$$

$$\tag{D.5}$$

$$cov[\hat{R}_k(\eta), \hat{R}_k(\eta + \tau)] = \frac{1}{2\pi} \int_{-\infty}^{\infty} [G_{11}(\omega)G_{22}(\omega) + G_{12}^2(\omega)e^{j\omega 2\eta}] \bullet e^{j\omega\tau} d\omega$$

The GCC trajectory over $[-T_m, T_m]$ is formed by filtering $\hat{R}_{12}$ with the time domain equivalent filter, $w(t)$, of the GCC weight $W(\omega)$. Let $t_1, \cdots, t_r$ be r distinct times in $[-T_m, T_m]$. Then write, $\underline{R}^g$ the vector of r samples of $\underline{R}^g$

$$\underline{R}^g = \frac{1}{m} \sum_{k=1}^m \underline{R}_k^g \tag{D.6}$$

where

$$\underline{R}_k^g = [\hat{R}_{12}^k(\tau) * w(\tau)|_{\tau = t_1}, \cdots, \hat{R}_{12}^k(\tau) * w(\tau)|_{\tau = t_r}]^T \tag{D.7}$$

has mean

$$\mu^g = [R_{ss}(\tau - D) * w(\tau)|_{\tau = t_1}, \cdots, R_{ss}(\tau - D) * w(\tau)|_{\tau = t_r}]^T \tag{D.8}$$

and covariance

$$\sum = ((\sigma_{t_i, t_j})) \tag{D.9}$$

$$\sigma_{\eta, \eta + r} = \frac{1}{2\pi} \int_{-\infty}^{\infty} [G_{11}(\omega)G_{22}(\omega) + G_{12}^2(\omega)e^{j\omega 2\eta}]|W(\omega)|^2 e^{j\omega r} d\omega$$

where $\sigma_{\eta, \eta + r}$ is finite on $\mathbf{R}^2$. Now by a multi-dimensional version of the Berry-Esseen Theorem [8], if P_m is the probability law of $\sqrt{m} \sum^{-1/2}[\underline{R}^g - \mu^g]$ and G is the standard Gaussian law; there exists a constant $C(r)$ such that

$$sup|P_m - G| \leq C(r) \rho_3 \frac{1}{\sqrt{m}} \tag{D.10}$$

where

$$\rho_3 = E\{|\sum^{-1/2}(\underline{R}_1^g - \mu)|^3\} \tag{D.11}$$

and the *sup* is over the class of measurable convex subsets of r-dimensional space. Therefore, if $\rho_3 < \infty$, the finite dimensional distributions of $\underline{R}^g(\tau)$ are approximately multivariate Gaussian with covariance $\frac{1}{m} \sum$ and mean μ^g, for m sufficiently large.

APPENDIX E

Let $x(\tau)$ be a Gaussian a.s. continuous process. Denote by $N(t_0, t)$ the number of upcrossings of a level $-\epsilon$ by $x(\tau)$ over an interval $[t_0, t]$. Here conditions will be given such that the conditional rate function of N, $\mathbf{E}_c[N(t_0, t)]$, exists and expressions for the rate function and the intensity, $\lambda_c(\tau)$, will be displayed. The conditions to be given will also be sufficient for the validity of Eqs. (22) and (28) with $\lambda = \lambda_c$.

The central result is the following proposition.

Proposition E.1

Let $x(\tau)$ be an a.s. continuous Gaussian process with mean $\mu(\tau)$ and covariance $\sigma(\tau, v)$. Let $r_{ij}(\tau, v) = \partial^{i+j}/\partial \tau^i \partial v^j \, \sigma(\tau, v)$ and assume the derivatives

$$\dot{\mu}(\tau), \; \dot{r}_{10}(\tau, \tau), \; r_{10}(\tau, t_0), \; r_{01}(t_0, \tau), \; r_{11}(\tau,\tau)$$

are continuous in τ. Then if the matrix

$$\Lambda = \begin{bmatrix} r_{00}(\tau, \tau) & r_{01}(\tau, \tau) & r_{00}(\tau, t_0) \\ r_{10}(\tau, \tau) & r_{11}(\tau, \tau) & r_{10}(\tau, t_0) \\ r_{00}(t_0, \tau) & r_{01}(t_0, \tau) & r_{00}(t_0, t_0) \end{bmatrix} \tag{E.1}$$

is positive definite the following holds:

The expected number of upcrossings of $-\epsilon$ by $x(\tau)$ on $[t_0, t]$ given $x(t_0) < -\epsilon$ is

$$E_c[N(t_0, \tau)] = \int_{t_0}^{t} \lambda_c(\tau) d\tau \tag{E.2}$$

where

$$\lambda_c(\tau) = K(\tau) \int_0^\infty y \, \Phi(ay + \theta) \exp\{-\frac{1}{2}(d_3 y + \theta)^2\} dy \tag{E.3}$$

and

$$\tag{E.4}$$

$$K(\tau) = \frac{1}{2\pi\sqrt{c}\,|\Lambda|^{1/2}} \quad \frac{1}{\Phi\left[\frac{-\epsilon - \mu(t_0)}{\sqrt{\sigma_{t_0 t_0}}}\right]} \quad \exp\left[-\frac{1}{2}\left(-\epsilon - \mu(\tau)\right)^2 \left(d_1 - \frac{d_2^2}{d_3}\right)\right]$$

$$a = \frac{b_2}{\sqrt{c}} \qquad (E.5)$$

$$\theta = \left[-\epsilon - \mu(t_0) + \frac{1}{c}\left[b_1(x_0 - \mu(\tau)) - b_2\eta(\tau)\right]\right]\sqrt{c} \qquad (E.6)$$

$$\xi = -\eta(\tau) + \frac{d_2}{d_3}\left(-\epsilon - \mu(\tau)\right) \qquad (E.7)$$

where in (E.4) - (E.7) the elements of Λ^{-1} are implicitly defined

$$\Lambda^{-1} = \begin{bmatrix} A & b \\ b^T & c \end{bmatrix} \qquad (E.8)$$

$$\begin{bmatrix} d_1 & d_2 \\ d_2 & d_3 \end{bmatrix} = A - \frac{1}{c}bb^T \qquad (E.9)$$

$$\begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = b \qquad (E.10)$$

Although the proof of the above proposition is straightforward, the details are tedious and will not be given here. The essential element is to show that the conditional joint density function, $g_{\tau,h}(y,z)$, of $x(\tau)$ and $\frac{x(\tau+h) - x(\tau)}{h}$ given $x(t_0) < -\epsilon$ satisfies the Leadbetter conditions given in Chapter 4, Section 1. Then Leadbetter's results in [12] imply that (E.2) holds for an intensity, $\lambda_c(\tau)$, equal to the infinite integral $\int_0^\infty z p_\tau(-\epsilon,z)dz$, where the conditional joint density $p_\tau(-\epsilon,z) = \lim_{h\to 0} g_{\tau,h}(-\epsilon, z)$. The specific form of the intensity $\lambda_c(\tau)$ is then obtained by manipulation of

$$\int_0^\infty z \lim_{h\to 0} g_{\tau,h}(-\epsilon, z)dz \qquad (E.11)$$

into the calculable form of (E.4).

APPENDIX F

Here moment conditions on the spectra of the signals and noises $s(t)$ and $n_1(t)$, $n_2(t)$ will be given to ensure the continuity conditions on the covariance function of R^g, stated in Proposition E.1. In the following, assume $R^g(\tau)$ has mean $\mu(\tau)$ covariance $\sigma(\tau, v)$ and let $r_{ij}(\tau, v)$ be defined as in Proposition E.1.

Lemma F.1

Assume $R^g(\tau)$ obeys a Gaussian probability law. Let the following four moment conditions hold. Then the derivatives $\dot{\mu}(\tau)$, $r_{10}(\tau, \sigma)$, $r_{01}(\tau, \sigma)$, $r_{11}(\tau, \tau)$ are continous in their arguments.

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} \omega^2 G_{ss}(\omega)d\omega \;<\; \infty \tag{F.1}$$

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} |\omega|^3 \, G_{ss}^2(\omega)d\omega \;<\; \infty \tag{F.2}$$

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} \omega^2 G_{11}(\omega)G_{22}(\omega)d\omega \;<\; \infty \tag{F.3}$$

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} |\omega|^3 G_{11}(\omega)G_{22}(\omega)d\omega \;<\; \infty \tag{F.4}$$

Proof

The existence of the derivatives of $\dot{\mu}$, r_{10}, r_{01} and r_{11} is sufficient for continuity. From the unbiasedness of the cross- correlation estimate

$$\mu(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} G_{ss}(\omega) W(\omega)e^{j\omega(\tau-D_0)}d\omega \tag{F.5}$$

and its second derivative

$$\mu''(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} -\omega^2 G_{ss}(\omega) W(\omega)e^{j\omega(\tau-D_0)}d\omega \tag{F.6}$$

Now since $W(\omega)$ is a finite energy filter, we can take $W(\omega) = 1$, in (F.6) and obtain the bound

$$|\dot{\mu}(\tau)| \leq \frac{1}{2\pi} \int\limits_{-\infty}^{\infty} \omega^2 G_{\bullet\bullet}(\omega)d\omega \qquad (F.7)$$

From Chapter 3 Appendix C

$$r_{00}(\sigma, v) = \frac{1}{2\pi T} \int\limits_{-\infty}^{\infty} G_{11}(\omega)G_{22}(\omega)e^{j\omega(v-\sigma)} + G_{\bullet\bullet}^2(\omega)e^{j\omega(v+\sigma-2D_0)}d\omega \quad (F.8)$$

Note from (F.8)

$$(F.9)$$

$$r_{ij}(\sigma, v) = \frac{1}{2\pi T} \int\limits_{-\infty}^{\infty} (j\omega)^{i+j} \left[(-1)^j G_{11}(\omega)G_{22}(\omega)e^{j\omega(v-\sigma)} + G_{\bullet\bullet}^2(\omega)e^{j\omega(v+\sigma-2D_0)}\right] d\omega$$

Hence

$$|r_{20}(\sigma, v)| \leq \frac{1}{2\pi T} \int\limits_{-\infty}^{\infty} \omega^2 \left[G_{11}(\omega)G_{22}(\omega) + G_{\bullet\bullet}^2(\omega)\right] d\omega \qquad (F.10)$$

which is bounded by (F.4) (recall $G_{11}G_{22} = G_{\bullet\bullet}^2 + G_{\bullet\bullet}[G_{n1} + G_{n2}] + G_{n1}G_{n2}$) clearly the right side of (F.10) also bounds $|r_{02}(\sigma, v)|$.

Finally, from (F.9)

$$\left|\frac{\partial}{\partial\tau} r_{11}(\tau, \tau)\right| \leq \frac{1}{2\pi T} \int\limits_{-\infty}^{\infty} |\omega|^3 \left[G_{11}(\omega)G_{22}(\omega) + G_{\bullet\bullet}^2(\omega)\right] d\omega \qquad (F.11)$$

or the boundedness of (F.3) and (F.5) guarantees (F.11) is bounded.

APPENDIX G

In this appendix an approximation to the conditioned intensity function $\lambda_c(\tau)$ given in Eq. (50) is derived. It is specifically assumed that the signal and noises possess flat bandlimited spectra with one sided bandwidth $\frac{B}{2}$ and spectral magnitudes S/B, N/B respectively.

First we need the quantities contained in the matrix Λ of Proposition E.1. In Chapter 3, Appendix C the covariance function, $\sigma_{\tau, \sigma}$ of $R^g(\tau)$ was derived. For coherent correlation processing the result is expressed in terms of the obser-

vation time T, the auto-spectrum and cross-spectrum of the observations, $G_{ij}(w)$, i,j = 1,2

$$\sigma_{v,s} = \frac{1}{2\pi T} \int_{-\infty}^{\infty} G_{11}(\omega)G_{22}(\omega)e^{j\omega(v-s)} + G_{12}^2(\omega)e^{j\omega(v+s)}d\omega \qquad (G.1)$$

The ambiguity process $\Delta R^g(\tau) = R^g(\tau) - R^g(D_0)$ has covariance

$$r_{00}(v,s) = \sigma_{v,s} + \sigma_{D_0,D_0} - \sigma_{D_0,s} - \sigma_{v,D_0} \qquad (G.2)$$

with first order mixed derivatives

$$r_{10}(v,s) = \frac{\partial}{\partial v}\sigma_{v,s} - \frac{\partial}{\partial v}\sigma_{s,D_0}, \qquad r_{11}(\tau,\tau) = \frac{\partial^2}{\partial v \partial s}\sigma_{v,s}\Big|_{v,s=\tau} \qquad (G.3)$$

For the flat signal and noise spectra of interest, and for $D_0 = 0$, specific expressions for the r_{ij} can be derived straightforwardly:

$$(G.4)$$

$$r_{00}(\tau,v) = \frac{(S+N)^2}{2\pi BT}h_{\tau,v}^{(1)} + \frac{S^2}{2\pi BT}h_{\tau,v}^{(2)}$$

$$r_{10}(\tau,v) = \frac{(S+N)^2}{4\pi T}\left[g\left(\frac{B}{2}(\tau-v)\right) - g\left(\frac{B}{2}\tau\right)\right] \qquad (G.5)$$
$$+ \frac{S^2}{4\pi T}\left[g\left(\frac{B}{2}(\tau+v)\right) - g\left(\frac{B}{2}\tau\right)\right]$$

and

$$r_{10}(\tau,\tau) = \frac{-(S+N)^2}{4\pi T}g\left(\frac{B}{2}\tau\right) \qquad (G.6)$$
$$+ \frac{S^2}{4\pi T}\left[g(B\tau) - g\left(\frac{B}{2}\tau\right)\right]$$

$$r_{11}(\tau,\tau) = \frac{(S+N)^2}{2\pi T \bullet 12}B + \frac{S^2}{2\pi T \bullet 4}B\,g'(B\tau) \qquad (G.7)$$

With respect to the expressions (G.4), (G.5), and (G.7) the following quantities have been defined

$$g(\tau) = \frac{\cos(\tau) - sinc(\tau)}{\tau} \qquad \text{(G.8)}$$

$$g'(x) = -sinc(x) - g(x)/x \qquad \text{(G.9)}$$

$$h_{\tau,\upsilon}^{(1)} = 1 + sinc\left[\frac{B}{2}(\tau - \upsilon)\right] - sinc(\frac{B}{2}\tau) - sinc(\frac{B}{2}\upsilon) \qquad \text{(G.10)}$$

$$h_{\tau,\upsilon}^{(2)} = 1 + sinc\left[\frac{B}{2}(\tau + \upsilon)\right] - sinc(\frac{B}{2}\tau) - sinc(\frac{B}{2}\upsilon) \qquad \text{(G.11)}$$

The mean of $\Delta R^g(\tau)$, $\mu(\tau)$, and its derivative, $\eta(\tau)$, have the expressions

$$\mu(\tau) = \frac{-S}{2\pi}\, h_\tau^{(3)} \qquad \text{(G.12)}$$

and

$$\eta(\tau) = \frac{S}{2\pi}\, \frac{B}{2}\, g(\frac{B}{2}\tau) \qquad \text{(G.13)}$$

where the additional quantity has been defined

$$h_\tau^{(3)} = 1 - sinc(\frac{B}{2}\tau)$$

The matrix Λ of interest is

$$\Lambda = \begin{bmatrix} r_{00}(\tau,\tau) & r_{01}(\tau,\tau) & r_{00}(\tau,-D_m) \\ r_{10}(\tau,\tau) & r_{11}(\tau,\tau) & r_{10}(\tau,-D_m) \\ r_{00}(-D_m,\tau) & r_{01}(-D_m,\tau) & r_{00}(-D_m,-D_m) \end{bmatrix}. \qquad \text{(G.14)}$$

It is straightforward, but tedious, to calculate the entries in Λ^{-1}, where Λ is as in Proposition E.1, Eq. (E.1). Implicitly define the matrix A and its entries a_i, the vector b and the b_i, and the scalar c as follows

$$\Lambda^{-1} = \begin{bmatrix} A & b \\ b^T & c \end{bmatrix} \qquad \text{(G.15)}$$

$$A = \begin{bmatrix} a_1 & a_2 \\ a_2 & a_3 \end{bmatrix} \qquad \text{(G.16)}$$

$$b = [b_1, b_2]^T \qquad \text{(G.17)}$$

The determinant can be calculated

$$|\Lambda| = \frac{1}{(2\pi T)^3} \; \frac{1}{B} \; \frac{1}{12} \; (S + N)^6 \, \Delta \tag{G.18}$$

where

$$\Delta = 3 + 2|\gamma_{12}|^2 + r(B\tau) \tag{G.19}$$

$|\gamma_{12}|$ is the magnitude coherency and $r(B\tau) = 0\left((B\tau)^{-2}\right)$ is a quantity which is much smaller than 1 if $B|\tau|, B|D_m - \tau| > 20\pi$.

For a given D_m, τ, define the quantities g_1, g_2, g_3 and g_4

$$g_1 = g\left[\frac{B}{2} (\tau - D_m)\right] - g\left(\frac{B}{2} \; \tau\right) \;, \quad g_2 = g\left[\frac{B}{2} (\tau + D_m)\right] - g\left(\frac{B}{2} \; \tau\right)$$

$$\tag{G.20}$$

$$g_3 = -g\left(\frac{B}{2} \; \tau\right) \;, \quad g_4 = g(B\,\tau) - g\left(\frac{B}{2} \; \tau\right) \;, \quad g' = g(B\,\tau)$$

With the above definitions, the entries in Λ^{-1} are

$$\tag{G.21}$$

$$a_1 = \frac{2\pi BT}{(S+N)^2\Delta} \left\{ [1 + 3|\gamma_{12}|^2 g'] \; [h^{(1)}_{-D_m,-D_m} + |\gamma_{12}|^2 h^{(2)}_{-D_m,-D_m}] - 3[g_1 + |\gamma_{12}|^2 g_2]^2 \right\}$$

$$\tag{G.22}$$

$$a_2 = \frac{-6\pi T}{(S+N)^2\Delta} \left\{ [g_3 + |\gamma_{12}|^2 g_4][h^{(1)}_{-D_m,-D_m} + |\gamma_{12}|^2 h^{(2)}_{-D_m,-D_m}] - [g_1 + |\gamma_{12}|^2 g_2][h^{(1)}_{r,-D_m} + |\gamma_{12}|^2 h^{(2)}_{r,-D_m}] \right\}$$

$$\tag{G.23}$$

$$a_3 = \frac{24\pi T}{(S+N)^2 B\Delta} \left\{ [h^{(1)}_{r,r} + |\gamma_{12}|^2 h^{(2)}_{r,r}] \; [h^{(1)}_{-D_m,-D_m} + |\gamma_{12}|^2 h^{(2)}_{-D_m,-D_m}] - [h^{(1)}_{r,-D_m} + |\gamma_{12}|^2 h^{(2)}_{r,-D_m}]^2 \right\}$$

$$\tag{G.24}$$

$$b_1 = \frac{2\pi BT}{(S+N)^2\Delta} \left\{ 3[g_3 + |\gamma_{12}|^2 g_4] \; [g_1 + |\gamma_{12}|^2 g_2] - [1 + 3|\gamma_{12}|^2 g'] \; [h^{(1)}_{r,-D_m} + |\gamma_{12}|^2 h^{(2)}_{r,-D_m}] \right\}$$

$$\tag{G.25}$$

$$b_2 = \frac{-2\pi T}{(S+N)^2\Delta} \left\{ 6[h^{(1)}_{r,-D_m} + |\gamma_{12}|^2 h^{(2)}_{r,r}] \; [g_1 + |\gamma_{12}|^2 g_2] - 6[g_3 + |\gamma_{12}|^2 g_4] \; [h^{(1)}_{r,-D_m} + |\gamma_{12}|^2 h^{(2)}_{r,-D_m}] \right\}$$

$$c = \frac{2\pi T}{(S+N)^2 \Delta} \left\{ [h_{r,r}^{(1)} + |\gamma_{12}|^2 h_{r,r}^{(2)}] \ [1 + 3|\gamma_{12}|^2 g'] - 3[g_3 + |\gamma_{12}|^2 g_4]^2 \right\} \quad \text{(G.26)}$$

The terms in Λ^{-1}, (G.21)-(G.26), are general expressions, valid for all τ such that $|B\tau| > 2\pi$ and $|B[\tau - D_m]| > 2\pi$. However, we are primarily interested in approximations to λ_c when τ is such that $O((B\tau)^{-2})$ and $O((B|D_m - \tau|)^{-2})$ terms can be neglected. For simplicity $O((B\tau)^{-2})$ will be used to denote both of the above O notations. With this in mind, up to $O((B\tau)^{-2})$, (G.21)-(G.26) are equivalent to the compact identities

$$a_1 = \frac{2\pi BT}{(S+N)^2} \frac{2 + |\gamma_{12}|^2}{3 + 2|\gamma_{12}|^2}$$

$$a_2 = \frac{\pi T}{(S+N)^2} \ 0\left(\frac{1}{B\tau} \right) \ , \quad a_3 = \frac{24\pi T}{(S+N)^2} \frac{1}{B} \quad \text{(G.27)}$$

$$b_1 = \frac{-2\pi BT}{(S+N)^2} \frac{1 + |\gamma_{12}|^2}{3 + 2|\gamma_{12}|^2} \ , \quad b_2 = \frac{-\pi T}{(S+N)^2} \ 0\left(\frac{1}{B\tau} \right) \quad \text{(G.28)}$$

$$c = \frac{2\pi BT}{(S+N)^2} \cdot \frac{2 + |\gamma_{12}|^2}{3 + 2|\gamma_{12}|^2} \quad \text{(G.29)}$$

For the above, the following approximations were used

$$h_{r,-D_m}^{(1)} = h_{r,-D_m}^{(2)} = 1 + 0(1/B\tau) \ \tau \neq -D_m$$

$$h_{t,t}^{(1)} = 2, \ h_{t,t}^{(2)} = 1, \ g_1 = g_2 = g_3 = g_4 = 0\left(\frac{1}{B\tau} \right)$$

where the above quantities are defined in (G.10), (G.11) and (G.20). From (G.27)-(G.29), we can compute the d_i making up the matrix

$$A - \frac{1}{c} b b^T = \begin{bmatrix} d_1 & d_2 \\ d_2 & d_3 \end{bmatrix}$$

in Proposition E.1. The results are

$$
\text{(G.30)}
$$

$$
d_1 = \frac{2\pi BT}{(S+N)^2} \; \frac{1}{2+|\gamma_{12}|^2} \quad , \quad d_2 = \frac{\pi T}{(S+N)^2} \; 0\!\left(1/B\tau\right) \quad , \quad d_3 = a_3
$$

We will also need the quantity

$$
\alpha = \frac{b_2}{\sqrt{c}\;\sqrt{d_3}} = 0\!\left(\frac{1}{B\tau}\right) \tag{G.31}
$$

which is a function of $B\tau$ but not of SNR.

Now from the equations for $\mu(\tau)$ and $\eta(\tau)$, (G.12) and (G.13), and the relations (G.27)-(G.29), we can derive expressions for θ and $\varsigma = \sqrt{d_3}\,\xi$ in (E.6) and (E.7) of Proposition E.1

$$
\theta = \left[-\mu(-D_m) - \frac{b_2}{c}\;\eta(\tau) - \frac{b_1}{c}\;\mu(\tau) \right] \sqrt{c} \tag{G.32}
$$

$$
\varsigma = \sqrt{d_3}\;\xi = \left[-\eta(\tau) - \frac{d_2}{d_3}\;\mu(\tau) \right] \sqrt{d_3} \tag{G.33}
$$

where $x_0 = -\epsilon$ is set to 0 in Proposition E.1. The specific results for the present case are, up to $0(1/B\tau)$

$$
\theta = \frac{|\gamma_{12}|\;\sqrt{\dfrac{BT}{2\pi}}}{[(2+|\gamma_{12}|^2)\,(3+2|\gamma_{12}|^2)]^{1/2}} \tag{G.34}
$$

$$
\varsigma = |\gamma_{12}|\;\sqrt{\frac{BT}{2\pi}}\; 0(1/B\tau) \tag{G.35}
$$

Note the similar dependencies of θ and ς on the quantity $|\gamma_{12}|\sqrt{BT/2\pi}$.

Recall the form of $K(\tau)$ in Eq. (E.4) with $x_0 = \epsilon = 0$

$$
K(\tau) = \frac{1}{2\pi\sqrt{c}\;|\Lambda|^{1/2}} \; \frac{1}{\Phi\!\left(\dfrac{\mu(-D_m)}{\sqrt{\sigma_{-D_m,\,-D_m}}}\right)} \; \exp\!\left[\mu(\tau)\;\sqrt{d_1 - \frac{d_2^2}{d_3}} \right] \tag{G.36}
$$

Here we will need $K(\tau)/d_3$. From the above it is straightforward to calculate, up to $O(1/BT)$ terms

$$(G.37)$$

$$K(\tau) = \frac{B}{\sqrt{12} \sqrt{2 + |\gamma_{12}|^2}} \cdot \frac{\phi\left[|\gamma_{12}| \sqrt{\frac{B}{2\pi}}\, T \, \frac{1}{\sqrt{2 + |\gamma_{12}|^2}} \right]}{\Phi\left[|\gamma_{12}| \sqrt{\frac{B}{2\pi}}\, T \right]} \frac{1}{\sqrt{2 + |\gamma_{12}|^2}}$$

We will investigate the form of the conditional intensity, $\lambda_c(\tau)$, under conditions of either large or small signal-to-noise ratio, $|\gamma_{12}| \sqrt{BT}$, or, alternately, large values of $B|\tau|$ and $B|\tau - D_m|$. We first consider the latter case, which corresponds to an investigation of $\lambda_c(\tau)$ outside of the high sidelobe region of the signal-auto-correlation function.

Make a change of variable, $v = \sqrt{d_3}\, y$, in the integral associated with the intensity function, Eq. (E.3)

$$\lambda_c(\tau) = K(\tau) \int_0^\infty y \Phi(ay + \theta) \exp\left(-d_3 (y + \xi)^2\right) \qquad (G.38)$$

This yields the equivalent expression

$$\lambda_c(\tau) = \sqrt{2\pi} K(\tau)/d_3 \, H(\alpha) \qquad (G.39)$$

where

$$H(\alpha) = \int_0^\infty y \Phi(\alpha y + \theta) \, \phi(y + \varsigma) dy \qquad (G.40)$$

Here α and ς are as defined in Eqs. (G.31) and (G.33), and ϕ is the standard Gaussian density function. The integral $H(\alpha)$ will be expanded in an asymptotic series about the parameter α, $\alpha(\tau) = 0\left(\frac{1}{B\tau} \right)$, defined in Eq. (G.31) to motivate the uniform approximation

$$H(\alpha) = \Phi(\theta) \, \phi(\varsigma) \qquad (G.41)$$

The next step is to show sufficient regularity of the integral $H(\alpha)$ in order to exchange derivative and integral operators. The first derivative of the integrand of $H(\alpha)$, (G.40), is uniformly continuous in α and has a bounded integral

$$\frac{\partial}{\partial \alpha}\ y\Phi(\alpha y + \theta)\ \phi(y + \varsigma) = y^2\phi(\alpha y + \theta)\ \phi(y + \varsigma) \qquad (G.42)$$

and

$$\int_0^\infty y^2\phi(\alpha y + \theta)\ \phi(y + \varsigma)dy \leq \frac{1}{\sqrt{2\pi}}\ \int_0^\infty y^2\phi(y + \varsigma)dy \qquad (G.43)$$
$$= -\varsigma\ \phi(\varsigma) + (\varsigma^2 + 1)\ [1 - \Phi(\varsigma)] = p_1(\varsigma)$$

p_1 is finite since for the Gaussian density ϕ [10]

$$\lim_{|\varsigma| \to \infty} \varsigma^n\ \phi(\varsigma) = 0 \qquad (G.44)$$

and

$$\lim_{|\varsigma| \to \infty} \varsigma^n\ [1 - \Phi(\varsigma)] = 0 \qquad (G.45)$$

Likewise, the second derivative of the integrand of $H(\alpha)$

$$\frac{\partial^2}{\partial \alpha^2}\ y\Phi(\alpha y + \theta)\ \phi(y + \varsigma)$$
$$= -y^2(\alpha y + \theta)\ \phi(\alpha y + \theta)\ \phi(y + \varsigma)$$

and

$$\left| \int_0^\infty y^2(\alpha y + \theta)\ \phi(\alpha y + \theta)\ \phi(y + \varsigma)\ dy \right| \qquad (G.46)$$
$$\leq \int_0^\infty (|\alpha|\ y^3 + |\theta|\ y^2)\ \phi(y + \varsigma)dy$$
$$= 3!\ |\alpha|\ \phi(\varsigma)\ D_{-4}(\varsigma) + |\theta|\ p_1(\varsigma) = p_2(\varsigma)$$

where $D_{-v}(z)$ is a parabolic cylinder function of order v [9] and $p_1(\varsigma)$ is as in (G.43). Note from [9]

$$\lim_{\varsigma \to \infty} \phi(\varsigma)\ D_{-4}(\varsigma) = 0 \qquad (G.47)$$

and since $\theta,\ \varsigma \sim |\gamma_{12}|\ \sqrt{BT/(4\pi)}$ (see Eqs. (G.34) and (G.35))

$$\lim_{|\theta|, |\varsigma| \to \infty} |\theta| \, p(\varsigma) = 0$$

The boundedness and continuity properties of the integral and the integrand respectively imply sufficient smoothness so that the Taylor remainder theorem can be invoked

$$H(\alpha) = H(0) + \alpha H'(0) + \alpha^2/2 \, H''(\overline{\alpha}) \qquad \text{(G.48)}$$

where $\overline{\alpha}$ is in the neighborhood of zero, and the differentiation and integration in H' and H'' can be interchanged. Performing the integrations involved in $H(0)$ and $H'(0)$ we obtain

$$\begin{aligned} H(\alpha) = \Phi(\theta) \left(\phi(\varsigma) - \varsigma[1 - \Phi(\varsigma)] \right) \\ + \alpha\phi(\theta) \, p_1(\varsigma) + 0(\alpha^2) \, p_2(\varsigma) \end{aligned} \qquad \text{(G.49)}$$

Note that since ς and α are $0(1/(B|\tau|))$, in the sense that both $1/(B|\tau|)$ and $1/(B|D_m - \tau|)$ dominate ς and α, Eq. (G.49) suggests the approximation given in Eq. (G.41). Investigation of the relative magnitudes of the various terms in Eq. (G.49) indicates that if both $B|\tau|$ and $B|\tau - D_m|$ exceed 20π then the approximation (G.41) is within about 5% of its actual value. Hence, the approximation (G.41) will be valid over the majority of the of the *a priori* interval for large BD_m for all SNR.

We next turn our attention to the case where the $0(B|\tau|)$ term $\alpha(\tau)$ is large, e.g. in the sidelobe region of the signal auto-correlation function. It can be shown that $\alpha = \alpha(\tau)$, Eq. (G.31), is bounded for all τ such that $B|\tau|, B|\tau - D_m| \geq 2\pi$. (Just use the relation $|g_i| \leq \dfrac{1}{\pi}$ in (G.25), (G.26) and (G.30)). Therefore, using (G.44), (G.45) and (G.47) in the expressions for p_1 and p_2, (G.43) and (G.46), $p_1(\varsigma)$ and $p_2(\varsigma)$ go to zero as ς and θ become large, which occurs when $|\gamma_{12}| \sqrt{BT} \to \infty$. However, if $|\gamma_{12}| \sqrt{BT}$ is not large, then the terms involving α, in Eq. (G.49), may be significant. Nonetheless, it is possible to show, by using

integration by parts of $H(\alpha)$ for $\alpha \neq 0$, that

$$(G.50)$$

$$H(\alpha) = \Phi(\theta)\,\phi(\varsigma) - \varsigma\big(1 - \Phi(\varsigma)\,\Phi(\theta)\big)$$

$$+ \frac{1}{1+\alpha^2}\left[\phi\left(\frac{\varsigma + \alpha\theta}{\sqrt{1+\alpha^2}}\right) - \frac{\varsigma + \alpha\theta}{\sqrt{1+\alpha^2}}\left[1 - \Phi\left(\frac{\varsigma + \alpha\theta}{\sqrt{1+\alpha^2}}\right)\right]\right]\phi\left(\frac{\theta - \alpha\varsigma}{\sqrt{1+\alpha^2}}\right)$$

$$+ \frac{1}{\alpha}\,R(\alpha)$$

where

$$|R(\alpha)| \le \frac{\varsigma}{\alpha}\left[\phi\left(\frac{\theta}{\alpha}\right) - \frac{\theta}{\alpha}\left[1 - \Phi\left(\frac{\theta}{\alpha}\right)\right]\right] \qquad (G.51)$$

Now as $|\gamma_{12}|\,\sqrt{BT}$, goes to zero for finite α, $R(\alpha) \to 0$, and from (G.50)

$$H(\alpha) \to \Phi(\theta)\,\phi(\varsigma) + \frac{1}{2\pi}\ \frac{1}{1+\alpha^2} \qquad (G.52)$$

The right hand side of (G.52) can be quite close to (G.41) in the high sidelobe regions where α may be large.

Combining (G.49) for small $0(B|\tau|)$, with (G.52) for large $0(B|\tau|)$, it follows that for the following approximation to the intensity

$$\lambda_c(\tau) = \frac{K(\tau)}{d_3}\ \Phi(\theta)\,\phi(\varsigma) \qquad (G.53)$$

the approximation error is small for the following cases $B|\tau|$ and $B|D_m - \tau|$ greater than 20π; $|\gamma_{12}|\,\sqrt{BT} \gg 1$; or both $|\gamma_{12}|\,\sqrt{BT} \ll 1$ and α^2 large. In the sequel we concentrate on the first of the three cases mentioned above.

The final form of the time independent intensity approximation used in Section VI is obtained by substituting the approximate expressions for θ and ς, (G.34) and (G.35), into (G.53). Using $\varsigma \sim 0$, we obtain

$$\lambda_c(\tau) = \frac{B_0}{\sqrt{12}}\ \frac{1}{\sqrt{2 + |\gamma_{12}|^2}}\ \frac{\Phi\big(\gamma'\,\sqrt{B_0 T}\,(3 + 2|\gamma_{12}|^2)^{-1/2}\big)}{\Phi\big(\gamma'\,\sqrt{B_0 T}\big)} \qquad (G.54)$$

$$\bullet \quad \sqrt{2\pi} \, \phi\!\left(\acute{\gamma} \, \sqrt{B_0 T}\right)$$

where we have defined the normalized signal bandwidth B_0 in Hz

$$B_0 = \frac{B}{2\pi} \tag{G.55}$$

and

$$\acute{\gamma} = \sqrt{\frac{|\gamma_{12}|^2}{2 + |\gamma_{12}|^2}}$$

For the example considered in Section VI, the function $\alpha(\tau)$ is plotted in Fig. 6.1. From Fig. 6.1 two things are clear. α is of small magnitude over the majority of the a priori region of delay $[-D_m, 2\pi/B]$ ($[-8, 0]$ in the figure). Also, α has a small average value over the regions where α itself is large. Therefore, the integral over the *a priori* region is approximately independent of linear terms in α. Consequently, the approximation, (G.54), can be said to be accurate to $O(\alpha^2)$.

APPENDIX H

Here we derive the identity

$$P(\omega_i \in [t, t + \Delta], \text{ some } i = 1, 2, \cdots, k) \tag{H.1}$$

$$= \lambda(t) \bullet \Delta \frac{\left[\int_I \lambda(s)\,ds\right]^{k-1}}{(k-1)!} \exp\{-\int_I \lambda(s)\,ds\}$$

where ω_i are discrete times within the interval $I = [t_0, t_f]$ at which a Poisson count process on I, $N^*(I)$, generates points. Here N^* is assumed to have the time continuous intensity λ, and Δ is an infinitesimal quantity.

The left hand side of (H.1) is equivalent to

$$\tag{H.2}$$

$$\sum_{\substack{m=1 \\ l=1}} P\!\left(\bigcap_{i=1}^{m-1} \{\omega_i \in [t_0, t]\}, \bigcap_{i=m}^{m+l-1} \{\omega_i \in [t, t + \Delta]\}, \bigcap_{i=m+l}^{k} \{\omega_i \in [t + \Delta, t_f]\}\right)$$

where (H.1) has merely been expressed as the union of the disjoint events: $m-1$ points preceed time t; l points inhabit $[t, t + \Delta]$, $k - (m+l)+1$ points follow time $t + \Delta$.

For sufficiently small Δ, the orderliness property of the Poisson process [15] guarantees that all but the following terms in (H.2) are negligible

$$\text{(H.3)}$$

$$\sum_{m=1}^{k} P\left[\bigcap_{i=1}^{m-1} \{\omega_i \in [t_0, t]\}, \{\omega_m \in [t, t + \Delta]\}, \bigcap_{i=m+1}^{k} \{\omega_m \in [t + \Delta, t_f]\} \right]$$

From the independent increment property of N^*, (H.3) is identically

$$\sum_{m=1}^{k} \frac{\left[\int_{t_0}^{t} \lambda(s)\,ds\right]^{m-1}}{(m-1)!} \int_{t}^{t+\Delta} \lambda(s)\,ds \frac{\left[\int_{t+\Delta}^{t_f} \lambda(s)\,ds\right]^{k-m}}{(k-m)!} \exp\{-\int_{t_0}^{t} \lambda(s)\,ds\} \quad \text{(H.4)}$$

Or use the continuity of $\lambda(s)$ to obtain from (H.4)

$$\lambda(t) \bullet \Delta \sum_{m=1}^{k} \frac{1}{(m-1)!(k-m)!} \left[\int_{t_0}^{t} \lambda(s)\,ds\right]^{m-1} \left[\int_{t}^{t_f} \lambda(s)\,ds\right]^{k-m} \quad \text{(H.5)}$$

Finally recognize the following to obtain (H.1): the sum in (H.5) is equivalent to

$$\frac{1}{(k-1)!} \sum_{m=0}^{k-1} \frac{(k-1)!}{m!(k-1-m)!} \left[\int_{t_0}^{t} \lambda(s)\,ds\right]^{m} \left[\int_{t}^{t_f} \lambda(s)\,ds\right]^{k-1-m} \quad \text{(H.6)}$$

$$= \frac{1}{(k-1)!} \left[\int_{t_0}^{t_f} \lambda(s)\,ds\right]^{k-1}$$

where the binomial theorem has been invoked.

REFERENCES

Chapter 5

1. G.C. Carter, "Time delay estimation for passive sonar signal processing," *IEEE Trans. on ASSP*, Vol. ASSP-29, No. 3, pp. 463-470, 1981.

2. V.H. MacDonald and P.M. Schultheiss, "Optimum passive bearing estimation in a spatially incoherent noise environment", *Journal of the Acoustical Society of America*, Vol. 46, No. 1, Part 1, pp. 37-43, 1969.

3. J.P. Ianniello, "Time delay estimation via cross-correlation in the presence of large estimation errors," *IEEE Trans. on ASSP*, Vol. ASSP-30, No. 6, pp. 998-1003, 1982.

4. S.K. Chow and P.M. Schultheiss, "Delay estimation using narrowband processes". *IEEE Trans. ASSP* Vol. ASSP-29, No. 3, pp. 478-484, 1981.

5. B.V. Hamon and E.J. Hannan, "Spectral estimation of time delay for dispersive and non-dispersive systems," *J. Royal Statist. Soc. Ser. C*, Vol. 23, pp. 134-142, 1974.

6. G.C. Carter, "The generalized correlation method for estimation of time delay," *IEEE Trans. ASSP* Vol. ASSP-24, No. 4, pp. 320-327, 1976.

7. G.C. Carter, A.H. Nuttall and P.G. Gable, "The smoothed coherence transform ," *Proc. IEEE* Vol. 61, No. 10, pp.1497-1498, 1973.

8. V.V. Sazonov, *Normal Approximation - Some Recent Advances*, Springer - Verlag, Berlin Heidelberg, 1981.

9. I.S. Gradshteyn and I.M. Ryzhik, *Table of Integrals, Series, and Products*, Academic Press, New York, 1980.

10. M. Abramowitz and I.A. Stegun, *Handbook of Mathematical Functions*, Dover, New York, 1972.

11. J. Abrahams, "A survey of recent progress on level crossing problems for random processes," *Office of Naval Research,* Technical Memorandum, June 21, 1983.

12. M.R. Leadbetter, "On crossings of levels and curves by a wide class of stochastic processes," *Ann. Math. Stat.,* Vol. 37, pp. 260-267, 1966.

13. L.E. Miller and J.S. Lee, "Bandpass correlator analysis for general input assumptions," *IEEE Trans. Inform. Theory,* Vol. IT-28, No. 6, pp. 973-976, 1982.

14. K. Scarbrough, R.J. Tremblay, and G.C. Carter, "Performance predictions for coherent and incoherent processing techniques of time delay estimation," *IEEE Trans. Acoust., Speech, Signal Processing,* Vol. ASSP-31, No. 5, pp. 1191-1196, 1983.

15. D. L. Snyder, *Random Point Processes,* Wiley - Interscience, New York, 1975.

16. M.R. Leadbetter, "Point processes generated by level crossings," P. A. W. Lewis, ed., Stochastic Point Processes, New York, Wiley, 1972.

17. H. Cramer and M.R. Leadbetter, *Stationary and Related Processes* Wiley, New York, 1967.

18. A. Weiss and E. Weinstein, "Fundamental limitations in passive time delay estimation - part I: narrowband systems" *IEEE Trans. Acoust., Speech, Signal Proc.,* Vol. ASSP-31, No. 2, pp. 472-485, 1983.

19. J.P. Ianniello, E. Weinstein, A. Weiss, "Comparison of the Ziv-Zakai lower bound on time delay estimation with correlator performance," *ICASSP-83 Proceedings,* Vol. 2 , pp. 875-878.

Lowpass Auto-correlation with Associated Parabola
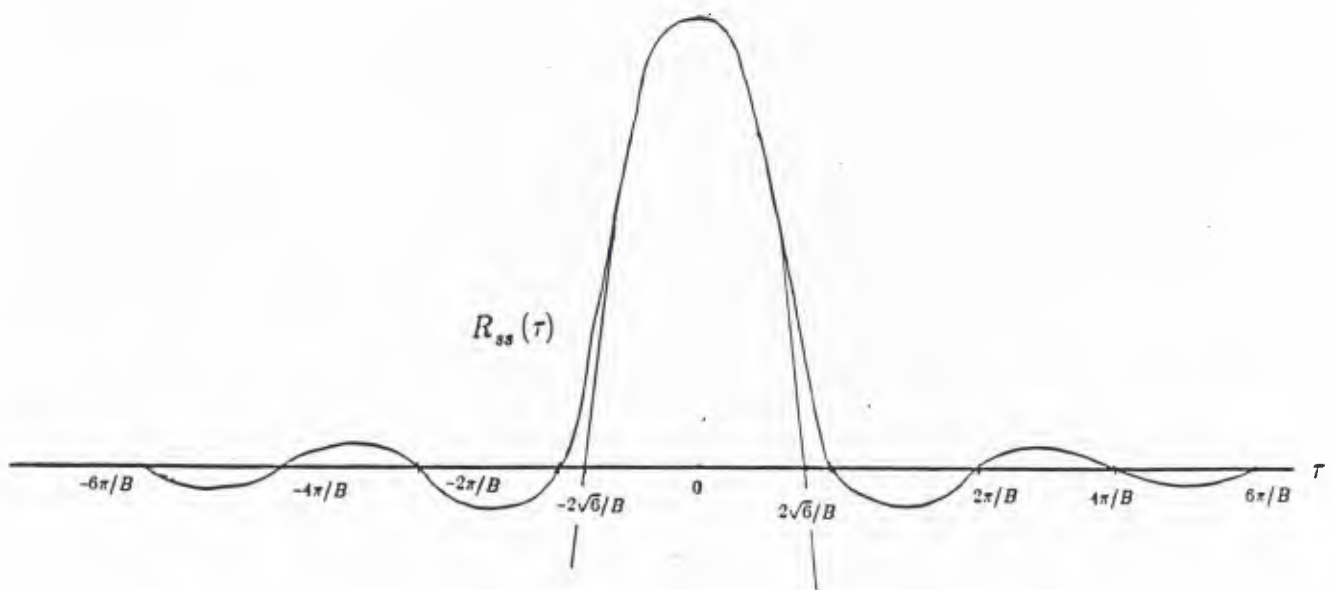
$$y(t) = R_{ss}(0) + \tau^2 R''(0)/2$$
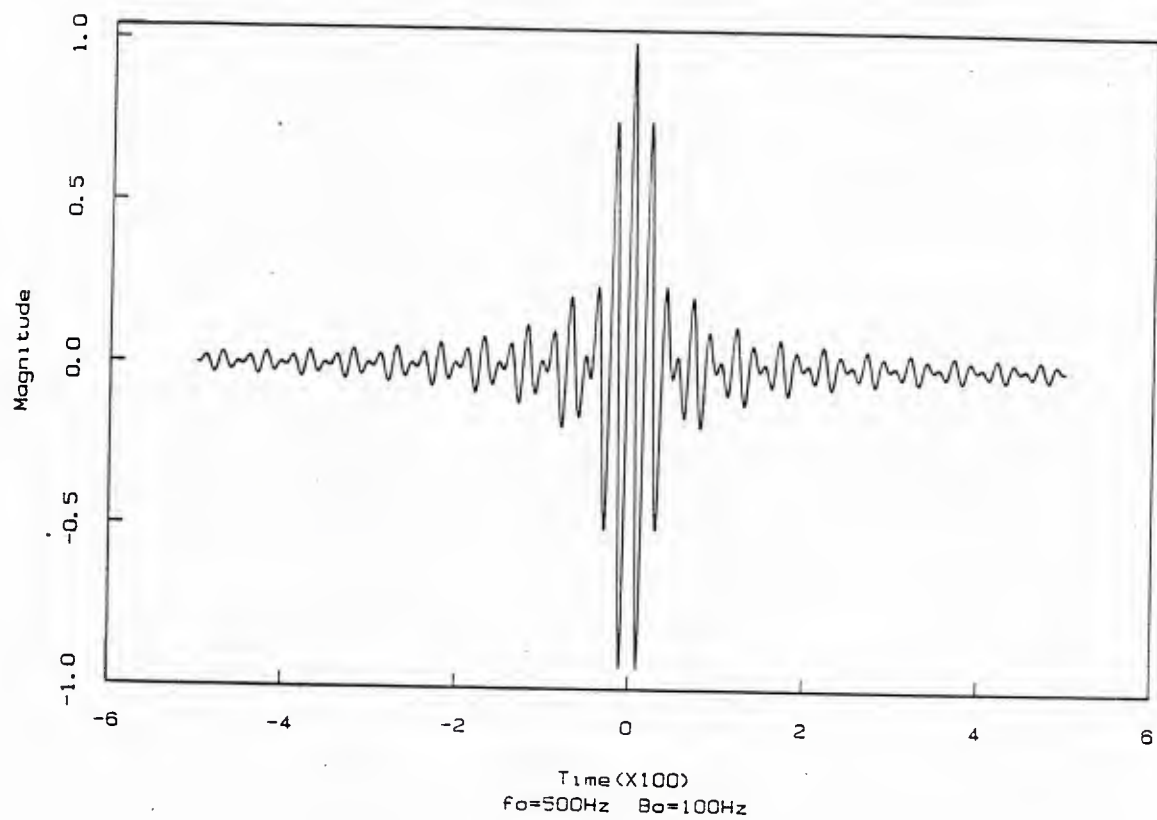
Fig. 2.1

Bandpass Signal Auto-correlation Function

Time (X100)
fo=500Hz Bo=100Hz

Fig. 2.2a

Lowpass Signal Auto-correlation Function

Time (X100)
Bo=100Hz
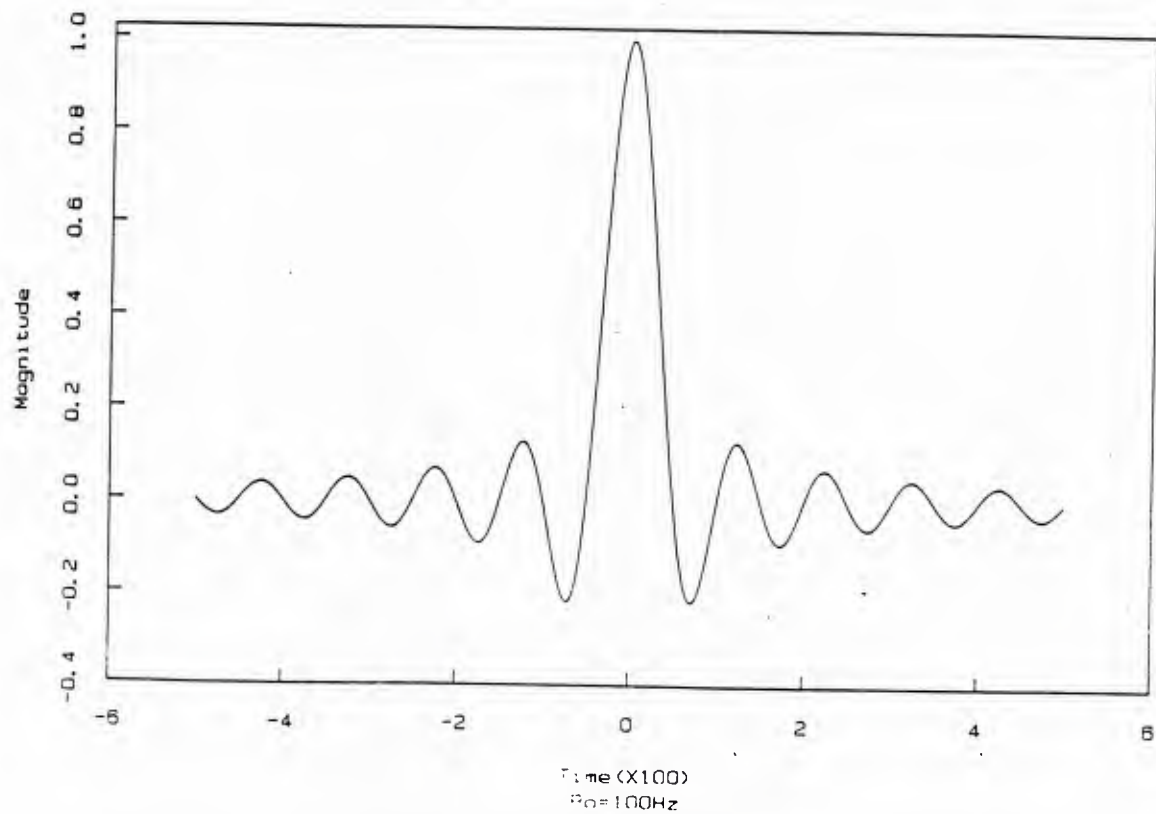
Fig. 2.2b

$\alpha(\tau)$ for Lowpass Signal Spectrum

Time (sec)
Bo=100Hz
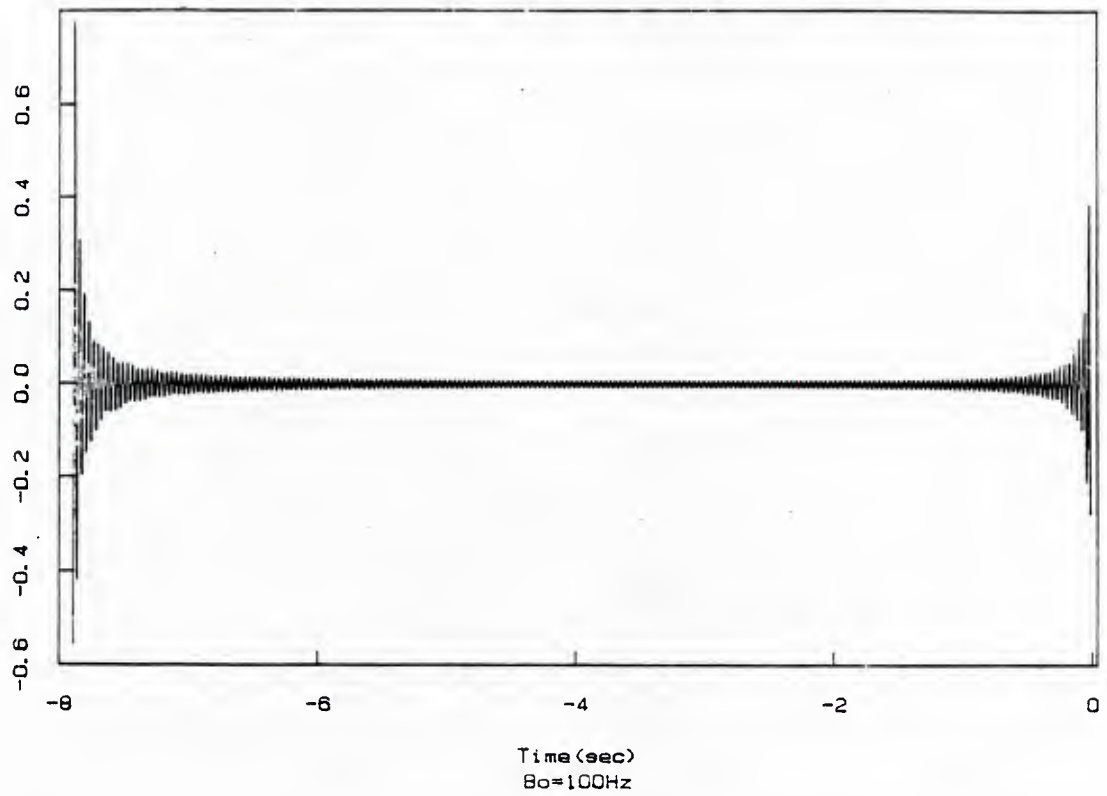
Fig. 6.1

Conditioned and Unconditioned Rate Functions

T=8.0sec.   Dm=0.125sec.
SNR (dB)

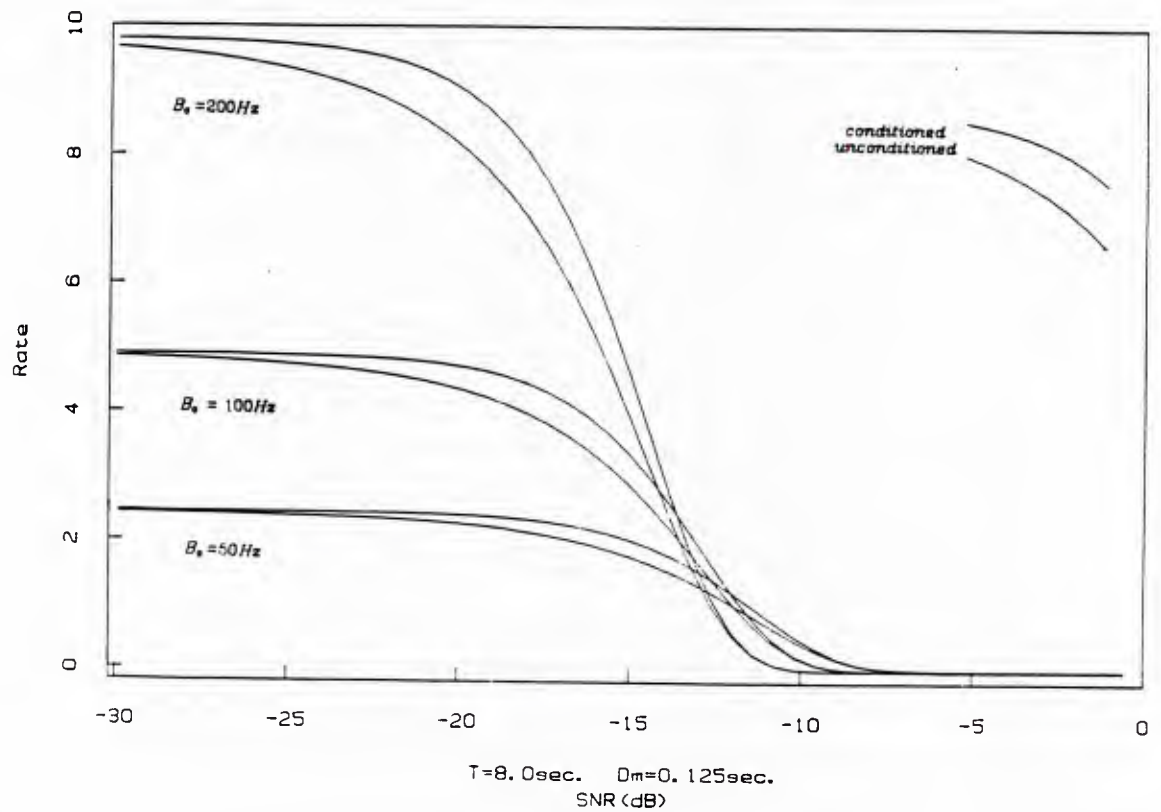Fig. 8.1

Intensity Surface for Lowpass Signal Spectrum
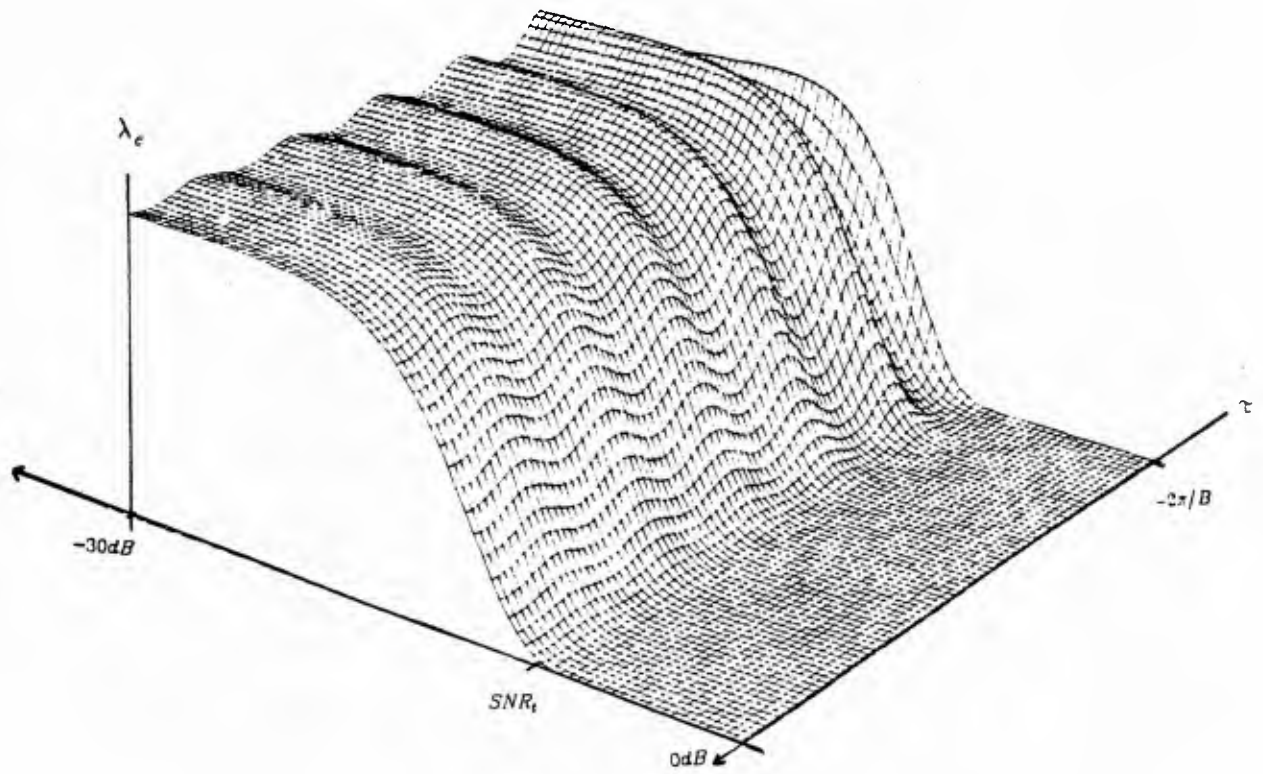
Fig. 8.2

Difference Between True and Uniform Approximation to Rate
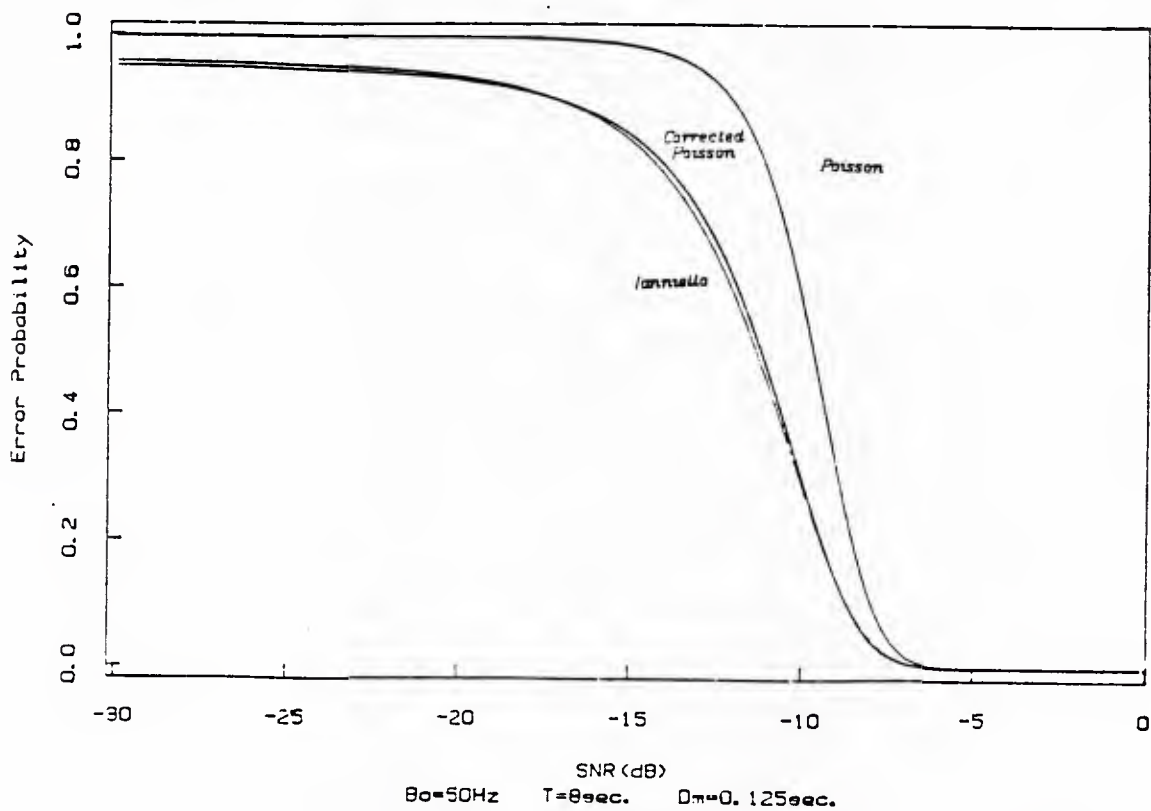
$D_m = 0.125sec.$
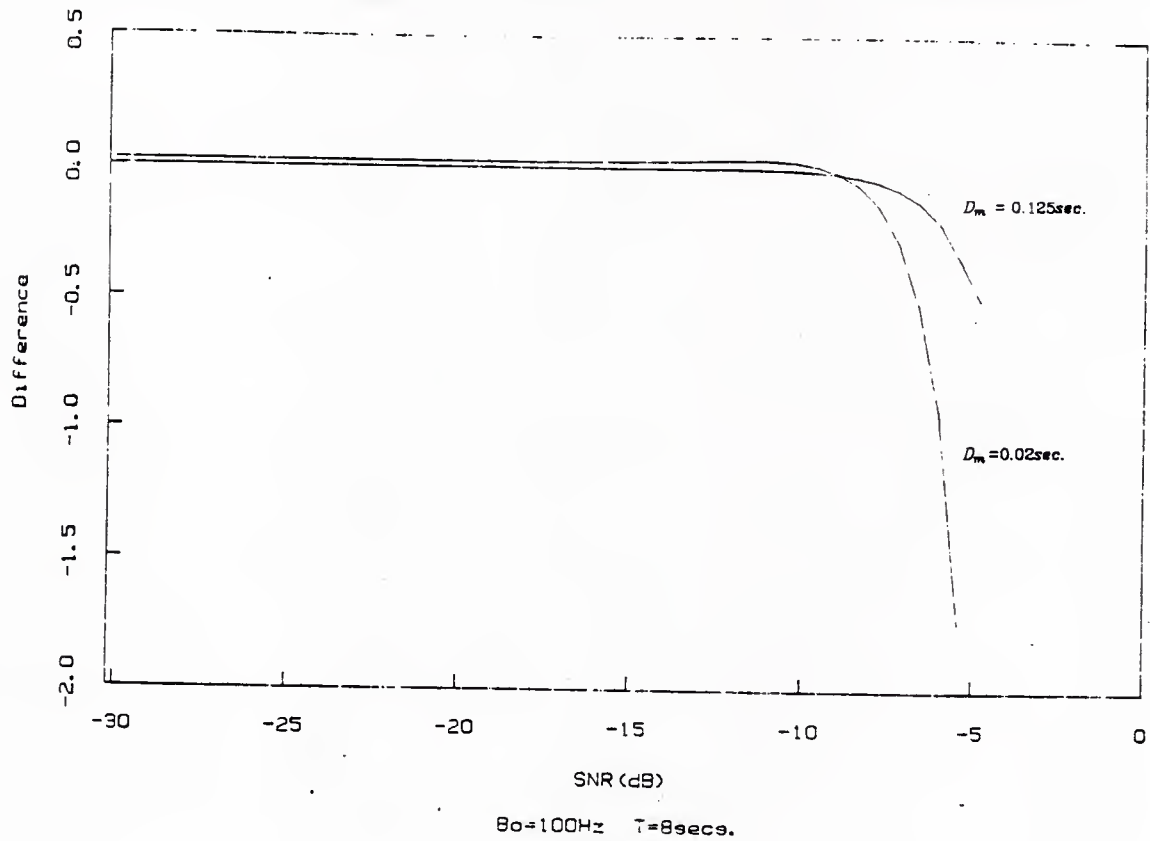
$D_m = 0.02sec.$

SNR (dB)

$Bo=100Hz$ $T=8secs.$

Fig. 8.3

Various Large Error Probability Approximations

Corrected Poisson

Poisson

Iannsello

SNR (dB)

$Bo=50Hz$ $T=8sec.$ $Dm=0.125sec.$

Fig. 8.4

Various Large Error Probability Approximations

SNR (dB)

Bo=100Hz   T=8sec.   Dm=0.125sec.

Fig. 8.5

Various Large Error Probability Approximations

SNR (dB)

Bo=600Hz   T=8sec.   Dm=0.125sec.
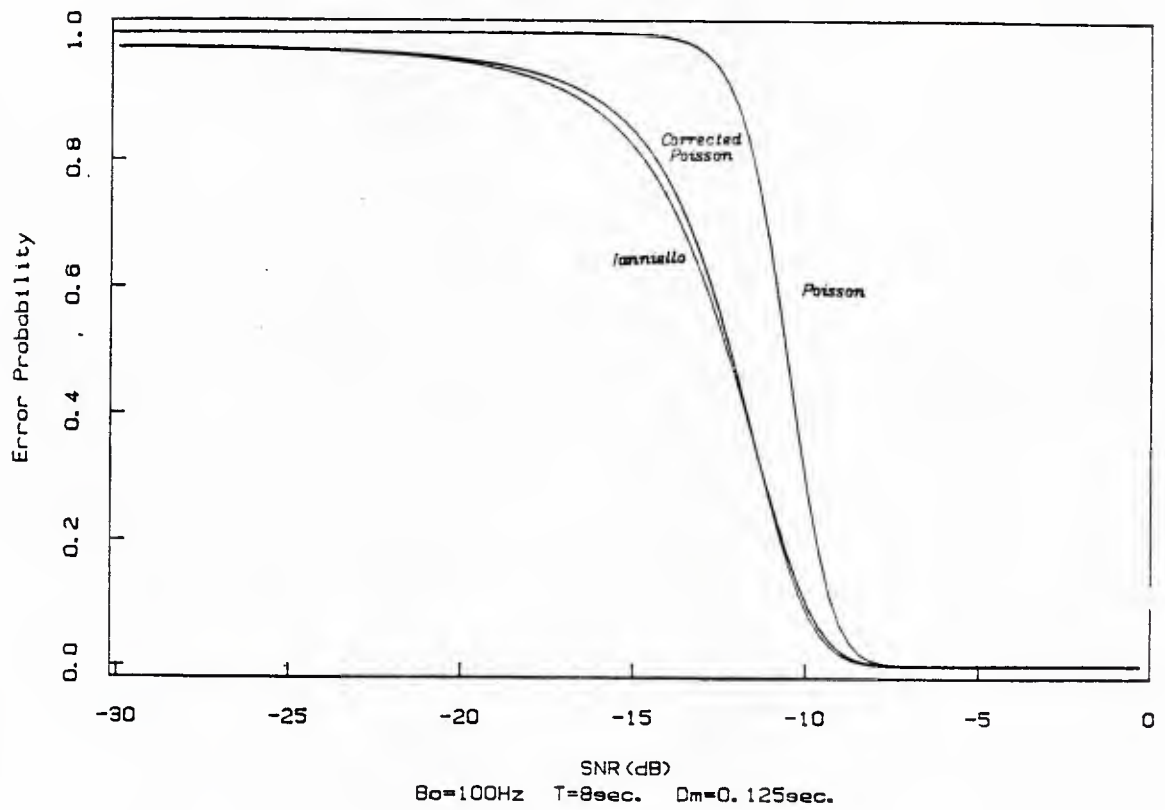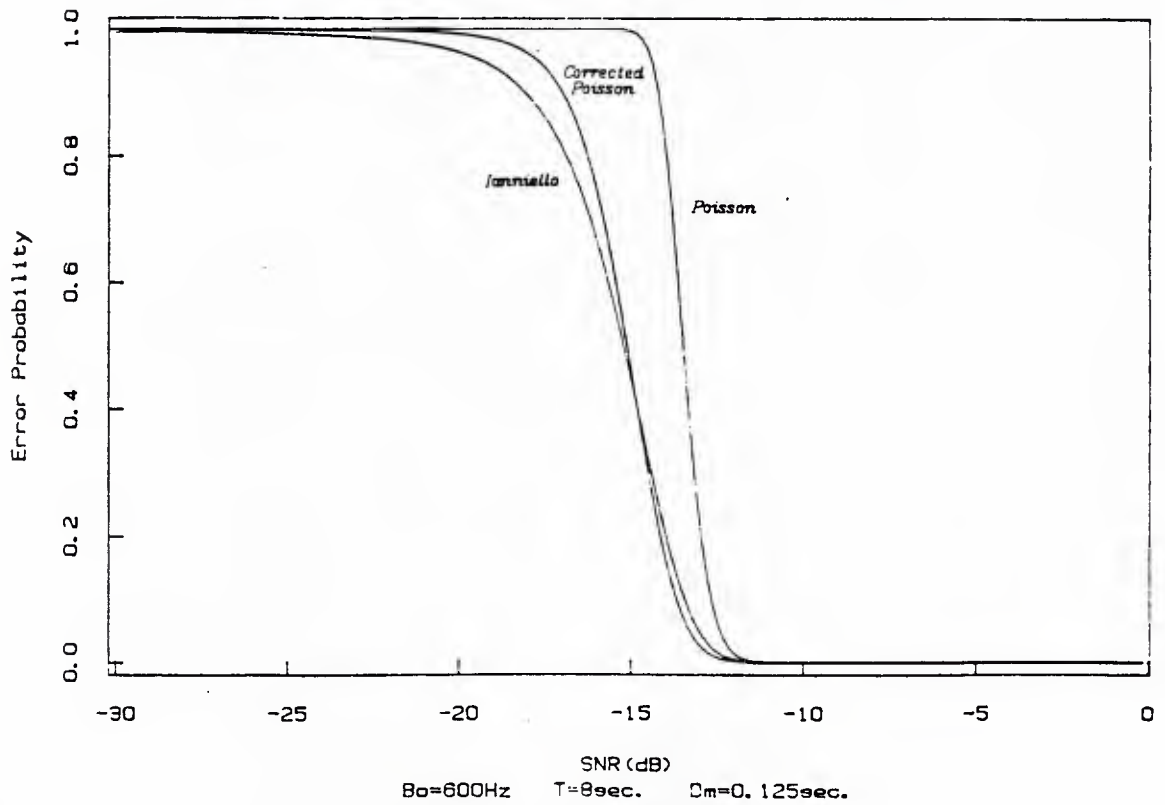
Fig. 8.6

Variance Approximations vs. Experimental Results

Bo=100Hz.   T=8secs.   Dm=0.125secs.
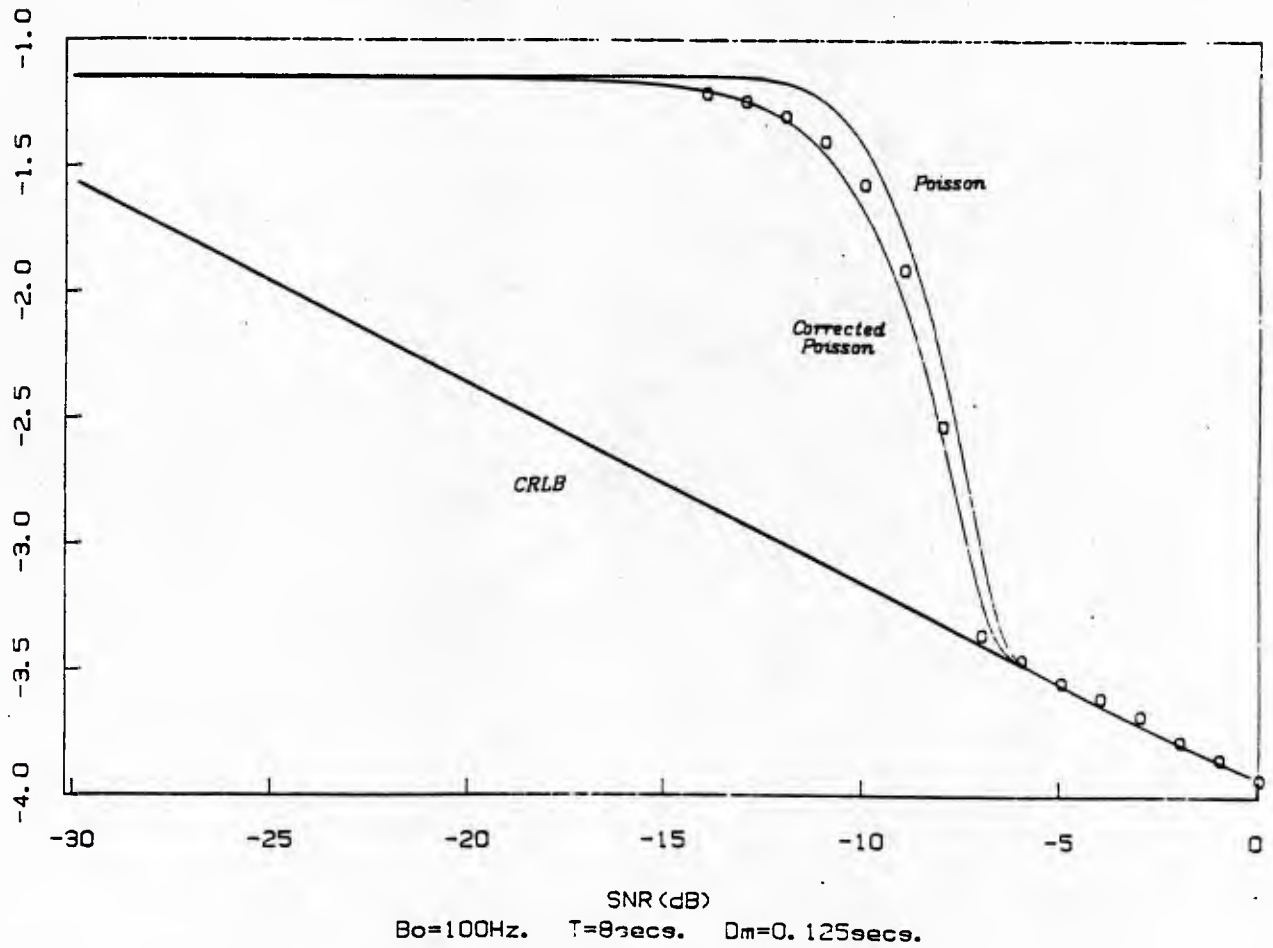
Fig. 8.7

Intensity Surface for Banpass Signal Spectrum
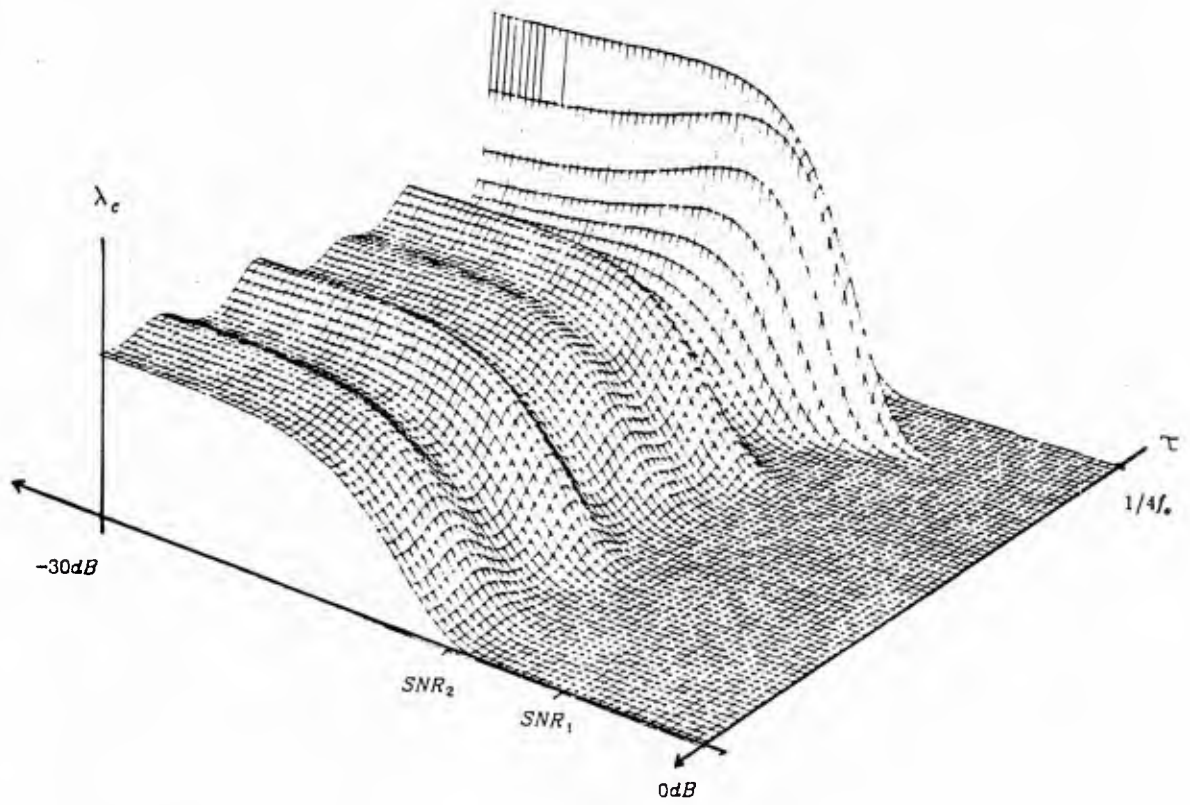
Fig. 8.8

SNR(dB) Thresholds Predicted by Ziv-Zakai Lower Bound

	SNR_3	SNR_2	SNR_1
$f_0 = 2000Hz$ $B_0 = 100Hz$ $T = 8.0sec.$ $D_m = 0.125sec.$	-16.6	-8.2	0.8
$f_0 = 2000Hz$ $B_0 = 500Hz$ $T = 8.0sec.$ $D_m = 0.125sec.$	-22.6	-19.7	-8.0
$f_0 = 500Hz$ $B_0 = 100Hz$ $T = 8.0sec.$ $D_m = 0.125sec.$	-19.7	-14.7	-8.7
$f_0 = 2000Hz$ $B_0 = 100Hz$ $T = 8.0sec.$ $D_m = 0.01sec.$	-10.9	-8.2	-0.8
$f_0 = 2000Hz$ $B_0 = 100Hz$ $T = 0.08sec.$ $D_m = 0.01sec.$	1.8	6.0	19.3

Fig. 8.9

Rate Function for High and Low Bandwidth

Bo = 500Hz

Bo = 100Hz

SNR (dB)

fo=2000Hz   Dm=0.125sec.   T=8.0sec.
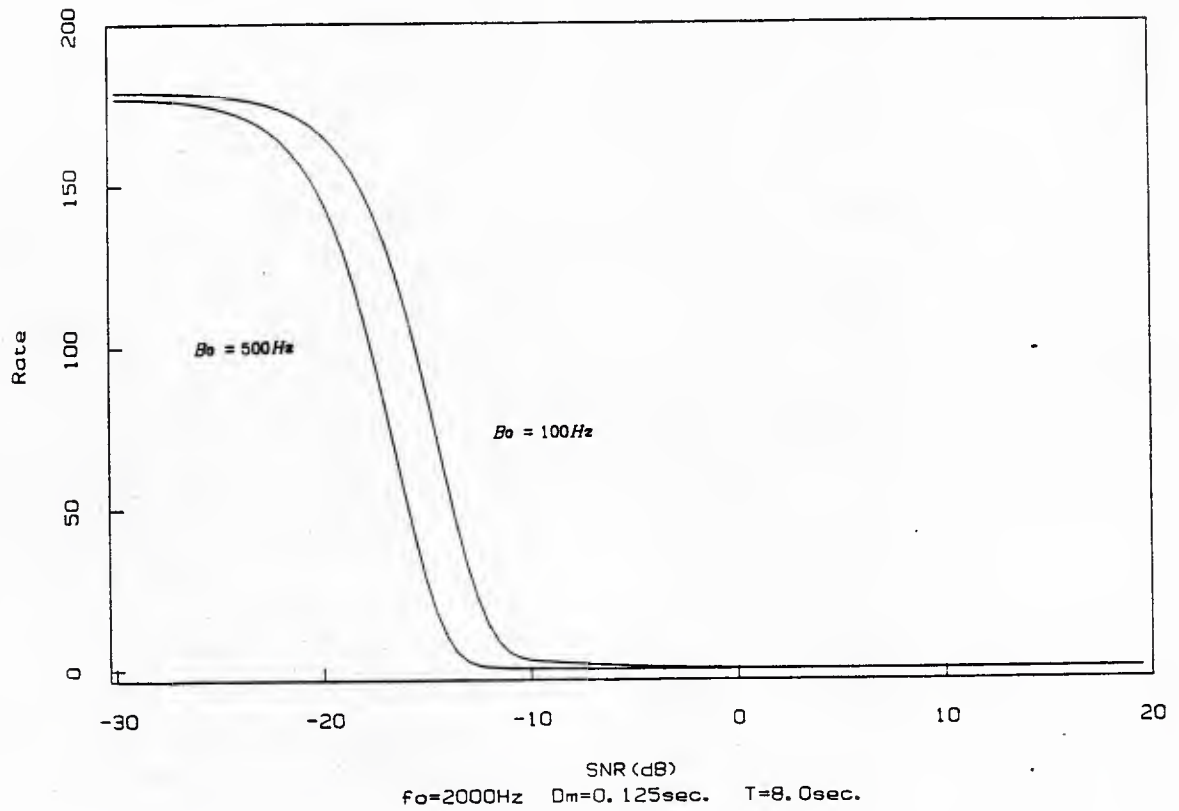
Fig. 8.10

Error Probability for High and Low Bandwidth

Bo = 500Hz

Bo = 100Hz

SNR (dB)

fo=2000Hz   Dm=0.125sec.   T=8.0sec.
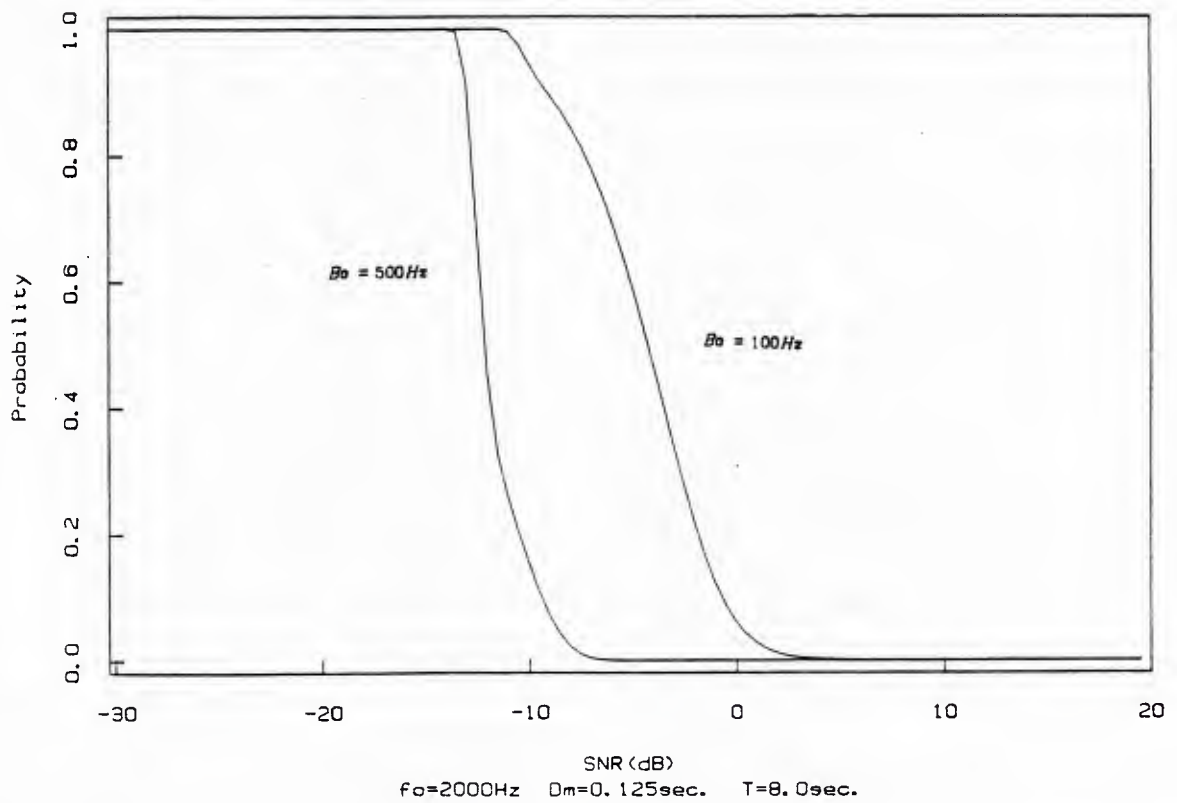
Fig. 8.11

Variance Approximations for High and Low Bandwidth

fo=2000Hz   Dm=0.125sec.   T=8.0sec.
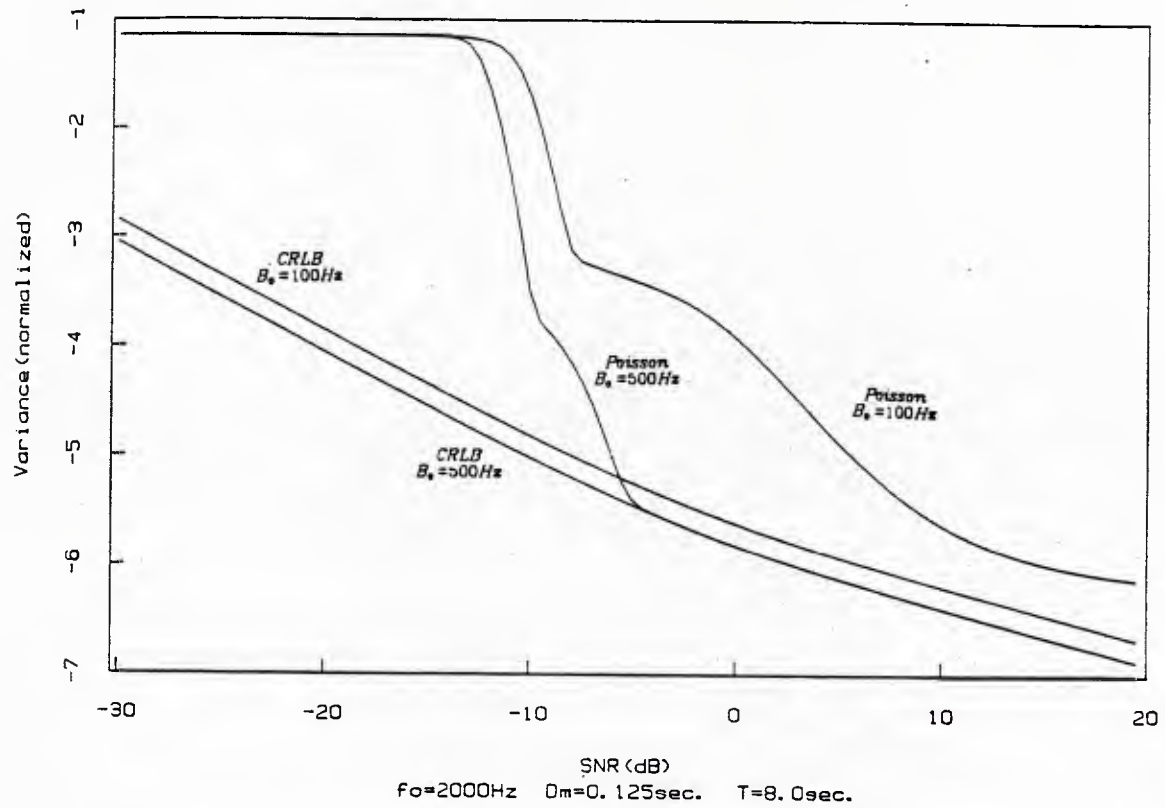
Fig. 8.12

Rate Function for Low and High   T

fo=2000Hz   B =100Hz   Dm=0.01sec.
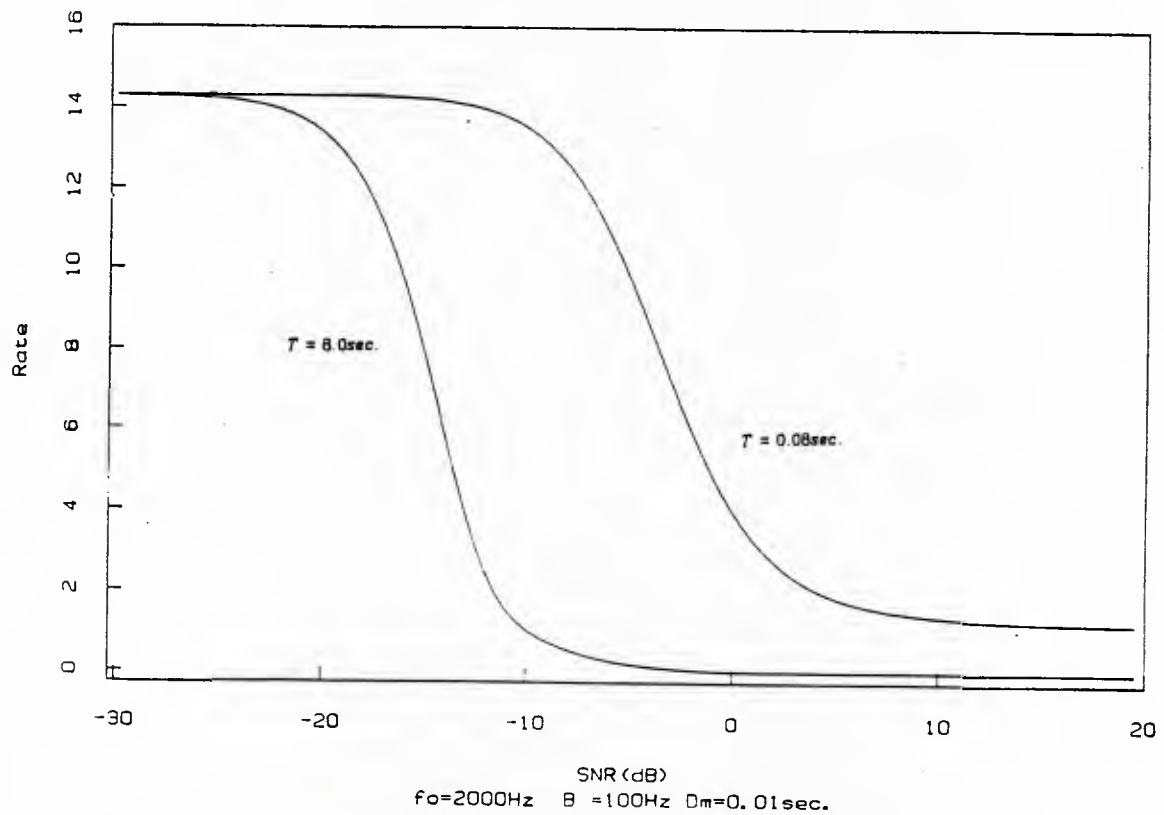
Fig. 8.13

Error Probability for High and Low T

SNR (dB)
Fo=2000Hz   Ro=100Hz   Dm=0.01sec.
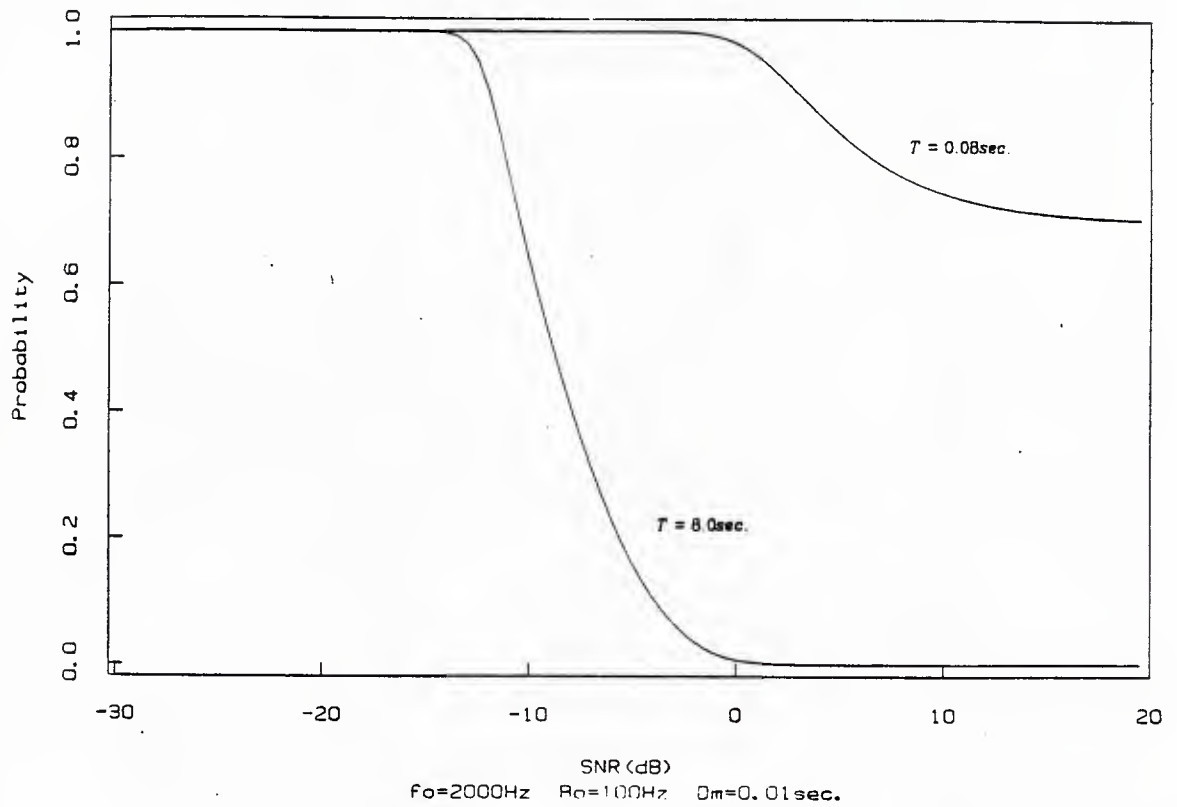
Fig. 8.14

Variance Approximations for High and Low T

SNR (dB)
Fo=2000Hz   Bo=100Hz   Dm=0.01sec.
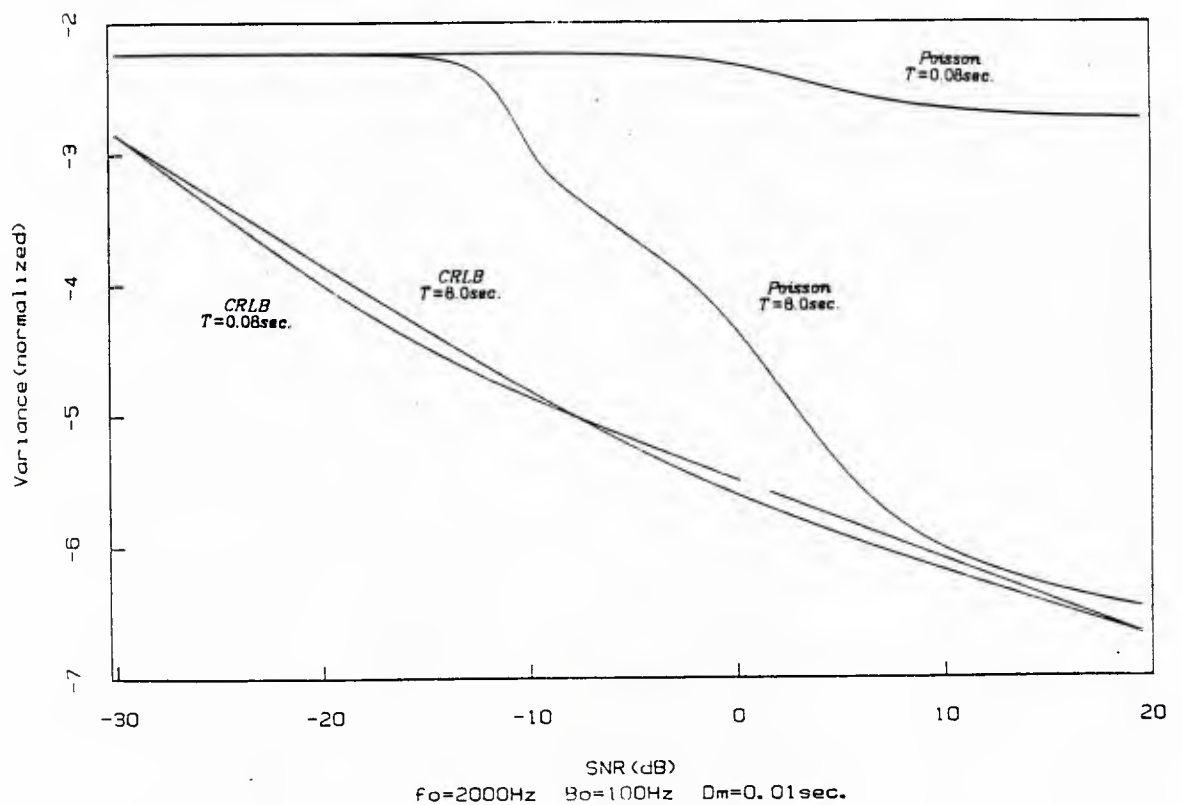
Fig. 8.15

Rate Function for High and Low fo

SNR (dB)

Bo=100Hz   Dm=0. 125sec.   T=8. 0sec.
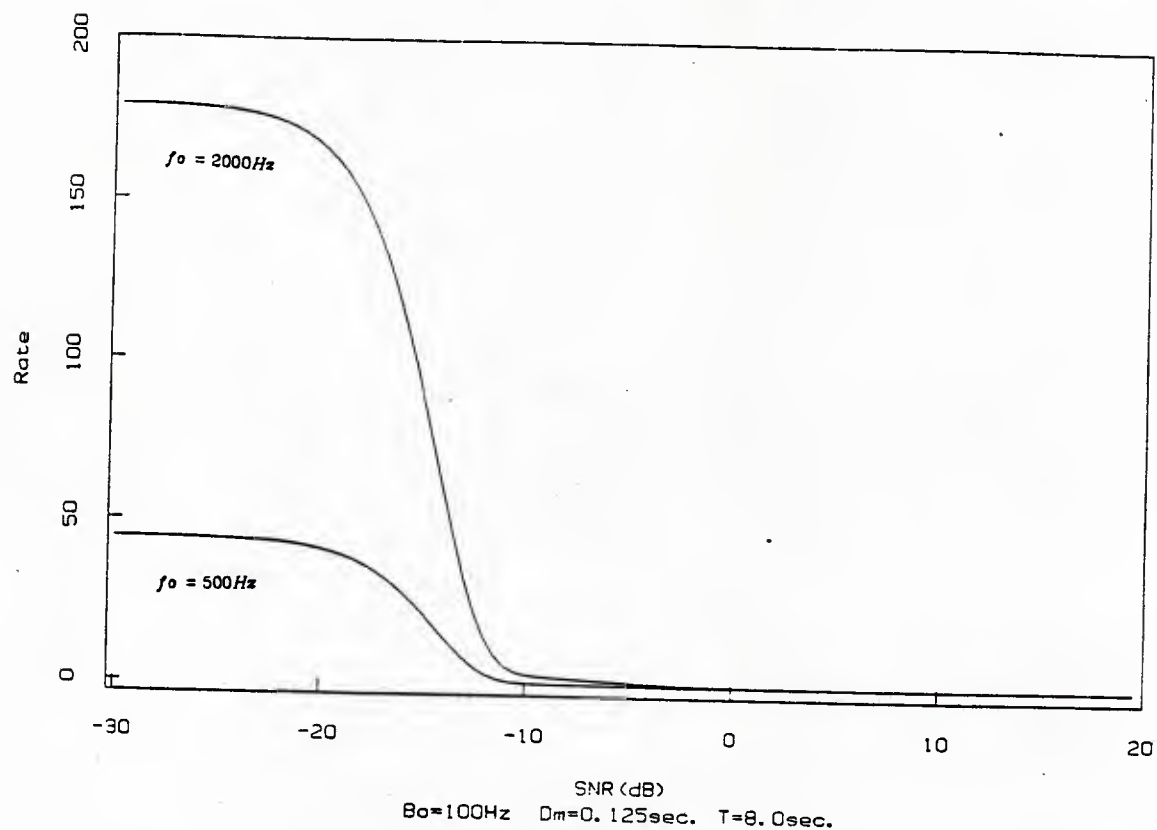
Fig. 8.16

Error Probability for High and Low fo

SNR (dB)

Bo=100Hz   Dm=0. 125sec.   T=8. 0sec.
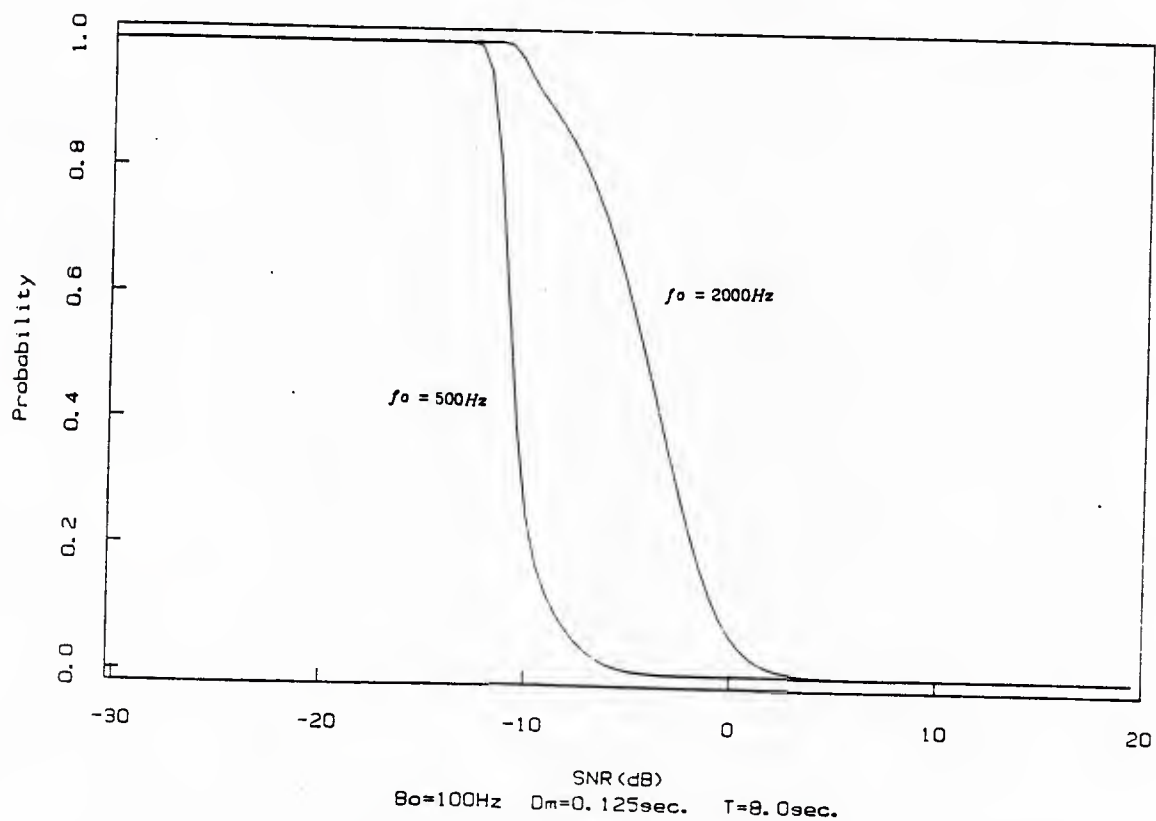
Fig. 8.17

Variance Approximations for High and Low fo
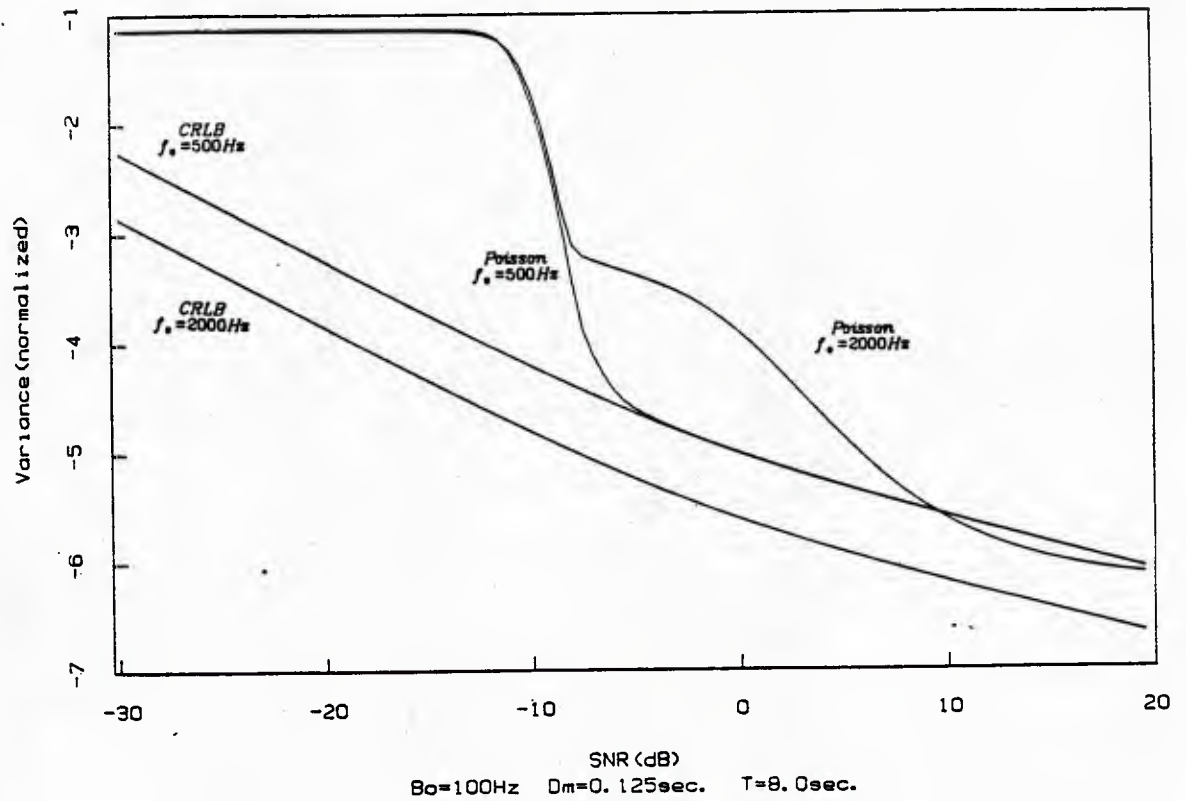
Fig. 8.18

Chapter 6
Additional Topics

I. Introduction

In this chapter two novel time delay estimators will be presented. Because of time and space limitations, a comprehensive analysis is not the objective here; rather it is to explore in a preliminary manner some interesting additional alternatives to the classical techniques. Hence, the presentation and development will be informal.

In Sections II, III and IV, we turn to time delay estimation for highly narrowband signals, that is when the cross-correlation function is oscillatory and ambiguity prone. A reinterpretation of the work of Chow and Schultheiss [4] on the Barankin lower bound reveals two facts: 1). our definition of SNR in Chapter 2, SNR_1, plays a major role in the determination of the lower bound on achievable estimation error , and 2). phase information contained in the observations becomes essentially useless at very low SNR. The above observations are used to motivate a hybrid estimator structure, a GCC which acts on both the observations and their envelopes. We then seek to optimize the structure by penalizing poor local (small error) performance while maintaining acceptable (low) sidelobe activity (large error). The form of the optimal hybrid processor is then displayed. This involves locally optimal GCC processing, combined with equalization via a tapped delay ine.

In Section V, a method is presented which transforms the inherently non-linear problem of estimating time delay into an (approximately) linear estimation problem. This is accomplished by operating on the cross-correlation trajectory with a linear, invertible, transformation, which demodulates the time delay directly onto the correlation domain. Optimal and sub-optimal procedures are then advanced as possible estimation schemes.

II. Background

The observations, over $[0, T]$, containing the time delay parameter, D, are as in chapter 3, Eq. (1). That is, it is assumed that two stationary Gaussian waveforms are available, each containing a common, but delayed in one case, Gaussian signal, $s(t)$, here of one-sided bandwidth B, buried in broadband mutually uncorrelated Gaussian noise processes , $n_1(t)$ and $n_2(t)$. For sufficiently large observation time-bandwidth product (BT) the Cramer-Rao Lower Bound (CRLB) on the variance of any estimate of D, $\hat{D}$ can be calculated [1], [2]. Specifically

$$J_{CRLB}(\hat{D}) = \frac{1}{2\pi} \int_{-\infty}^{\infty} w^2 \frac{|\gamma_{12}(w)|^2}{1 - |\gamma_{12}(w)|^2} \, dw \tag{1}$$

where $J_{CRLB}(\hat{D})$ is the inverse of the CRLB, or the Fisher information, and as usual, $|\gamma_{12}(w)|^2$ is the magnitude squared coherency defined in Chapter 2, Eq. (10). Since $|\gamma_{12}(w)|^2$ measures the importance of signal energy relative to noise energy at a given frequency w, $J_{CRLB}(\hat{D})$ can be interpreted as a measure of the frequency spread of usable signal energy for lowpass signals (See Fig. 2.2a).

Recall, at high signal-to-noise ratio (SNR) the performance of the GCC is well characterized by the local approximation to estimator variance, given by Chapter 3 Eq. (23). As discussed in Chapters 3 and 5 the local variance assumes its minimum value when the Hannan Thomson (H.T.) weight is used for $W(w)$ in the GCC. (For definitions of the GCC trajectory with weight $W(w)$ and the H.T. see Chapter 2, Eqs. (6) and (13)). Furthermore this minimum is identically the CRLB, the inverse of $J_{CRLB}(\hat{D})$. As was discussed in Chapter 2, the GCC performance measure SNR_1, ch. 2, Eq. (12), is maximized over all $W(w)$ by the $H.T.$. Consequently, it is clear that the $H.T.$ has a strong claim to optimality when BT and the signal-to-noise ratio are large.

However, SNR_1 is only reliable as a local (small error) performance measure since it applies equal penalty to all errors, large and small (Recall discussion in Section II of Chapter 2). In addition the CRLB on the error variance is unachievable below a certain SNR threshold determined by the center frequency to bandwidth ratio (w_o/B) and BT (for bandpass signals) [3], [4]. Therefore the H.T. can only be considered optimal in a restricted sense, which we will refer to as local optimality. In [3] a tighter lower bound on the error variance than the CRLB was derived, based on the Ziv-Zakai lower bound (ZZLB). The ZZLB based bound can be considered a composite bound made up of the following as the SNR varies from high to low values: the CRLB, a version of the ZZLB, the Barankin lower bound (BLB), and finally the lower bound given by uniformly distributed estimates throughout the *a priori* region of time delay. A typical example of the behavior of the composite bound is roughly reproduced from [3] in Fig. 2.1 for clarity.

Using intuition gained from an investigation of the specific form of the Barankin bound derived in [4], a more global measure of performance is proposed, i.e., a measure which is reliable for a greater range of SNR. This measure can then be optimized to yield an interesting processor structure.

III. Interpretation of the Barankin Bound

In its original form, as proposed by Barankin [5], the Barankin lower bound is a greatest lower bound on an unbiased estimator's error variance. In other words, there always exists an estimator which achieves the Barankin bound. However the form of the bound and the construction of the optimal estimator are difficult. The specific form of the Barankin lower bound used in the context of time delay estimation [4], the BLB, is a considerably weaker version of the original Barankin bound and is not acheivable in general. Nevertheless it is tighter

than the CRLB for sufficiently low SNR, by orders of magnitude in some cases, and as such gives more insight into the limiting factors affecting performance.

In [4] a general expression for the inverse of the Barankin bound, which we will refer to as $J_B(\hat{D})$, in the context of time delay estimation is given for narrowband bandpass signal spectra. The result in [4] is derived under the assumptions of multiple (M) sensors, white noises, large BT, and large center frequency-to-bandwidth ratio (w_o/B). This expression, [4, Eq. (20)], will be adapted to the present case of interest, i.e. $M=2$ and low SNR. In Eq. (20) of the reference, identify the ratio of signal and noise spectra as the magnitude coherency squared, $|\gamma_{12}(w)|^2$. Use the following low SNR approximation $(G(w) \alpha SNR)$: $\exp\{-\int \ln(1 - G(w))dw\} - 1 = \int G(w)dw$, to obtain an expression equivalent to Eq. (20) of [4]:

$$J_B(\hat{D}) = \frac{1}{2\pi} \int_{-\infty}^{\infty} (|w| - w_o)^2 \frac{|\gamma_{12}(w)|^2}{1 - |\gamma_{12}(w)|^2} \, dw \qquad (2)$$

In (2) we have transformed the summation in Eq. (20) of [4], indexed over a dense set within a segment of the positive real line, to an integral over the entire real axis, by using identical large BT assumptions as were used in [4] to obtain Eq. (23) of [4].

The important difference between $J_B(\hat{D})$ and $J_{CRLB}(\hat{D})$, (1) and (2), is that while $J_{CRLB}(\hat{D})$ measures the zero centered moment of the spectral quantity, $|\gamma_{12}(w)|^2/(1 - |\gamma_{12}(w)|^2)$, $J_B(\hat{D})$ measures the second moment centered about the center frequency w_o. Thus $J_B(\hat{D})$ is a more accurate indicator of the actual spread of signal energy over frequency than $J_{CRLB}(\hat{D})$ (See Fig. 2.1b). It is useful to investigate the form of $J_B(\hat{D})$ for highly narrowband signal spectra. Multiply out the square term in (2) and distribute the integral over the resultant three terms. Note that if $|\gamma_{12}(w)|^2$ is highly concentrated about the center frequency w_o then $|w|/w_o$ is approximately unity over the w-region where $|\gamma_{12}(w)|^2$ is

significantly non-zero. Hence we have as an approximation

$$J_B(\hat{D}) = \frac{1}{2\pi} \int_{-\infty}^{\infty} w^2 \frac{|\gamma_{12}(w)|^2}{1 - |\gamma_{12}(w)|^2} \, dw - w_o{}^2 \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{|\gamma_{12}(w)|^2}{1 - |\gamma_{12}(w)|^2} \, dw \quad (3)$$

One can now identify the difference in (3) in terms of the minimum local variance, Eq. (23) of Chapter 3, $var_L^o = \min\{var_L(\hat{D})\}$, i.e. the CRLB, and the maximum of SNR_1, Eq. (12) of Chapter 2, $SNR_1^o = \max\{SNR_1\}$, both achieved for the H.T. GCC

$$J_B(\hat{D}) = \frac{1}{var_L^o} - w_o{}^2 SNR_1^o \quad (4)$$

This interesting result implies that, while the local measures, var_L and SNR_1 are not individually sufficient to characterize the large error performance of the optimal estimator, a linear combination of these local criteria does parameterize the large error performance for sufficiently narrowband bandpass signal spectra. In the following section this fact is used to motivate a composite local performance criterion, $\rho = var_L + 1/w_o{}^2 \, SNR_1^{-1}$, where the normalization of SNR_1 is to make the two components of ρ of similar order of magnitude.

Observe that the BLB corresponds to the CRLB in the case where only the envelopes of the observed waveforms, x_1 and x_2 are available. Hence the high frequency components in the observations play no part in the BLB. In so far as the BLB is indicative of the optimal estimator's low SNR performance, this implies that the phase information in the bandpass signal, $s(t)$, is useless at very low SNR. It was suggested in [4] that at low SNR a near optimal time delay estimate can be obtained by passing the envelopes of the observed waveforms through a cross-correlator instead of the observables themselves. A more general implementation, using a GCC, will be referred to as the GCC envelope processor. Cross-correlation of the envelopes was motivated in [4] by consideration of the asymptotic form of $J_B(\hat{D})$, Eq. (20) of the reference, for the special case of flat

bandpass signal spectra. It is easily seen that any one of the GCC weights, $W(w)$, discussed in Chapter 2 is suitable for the GCC envelope processor in this particular case. In the general case, the local variance of the GCC envelope processor, $var_e(\hat{D})$, is straightforward to derive from the local variance expression of Schultheiss [6] (See Ch. 3 Eq. (23))

$$var_e(\hat{D}) = \frac{\frac{1}{2\pi T} \int_{-\infty}^{\infty} (|w| - w_o)^2 G_{11}(w) G_{22}(w)(1 - |\gamma_{12}(w)|^2)|W(w)|^2 dw}{\left[\frac{1}{2\pi} \int_{-\infty}^{\infty} (|w| - w_o)^2 |G_{12}(w)| W(w) dw\right]^2} \quad (5)$$

where, as in Chapter 3, the auto-spectra and cross-spectra of the observed waveforms have been denoted G_{11}, G_{22} and G_{12} respectively. The local variance expression of Eq. (5) is minimized by $W_{HT}(|w|-w_o)$, a frequency shifted version of the optimal H.T. weight given in Chapter 2 Eq. (13), and repeated here

$$W_{HT}(w) = \frac{1}{|G_{12}(w)|} \frac{|\gamma_{12}(w)|^2}{1 - |\gamma_{12}(w)|^2} \quad (6)$$

Furthermore the minimum of $var_e(\hat{D})$, taken at $W(w) = W_{HT}(|w|-w_o)$, is identically the BLB, the inverse of $J_B(\hat{D})$ in (2). Hence the H.T. envelope processor is locally optimal for narrowband bandpass signals at low SNR.

The expression, (2), for the inverse of the BLB is only valid for very low SNR. Near threshold, where the BLB is exceeded first by the Ziv-Zakai and finally by the CRLB, exclusive processing of only the envelopes of the observations entails a loss in important phase information. In the next section a more sophisticated processor is proposed, which is based on the optimization of a hybrid GCC-GCC-envelope-processor structure. The optimization is performed with respect to the local variance of this hybrid. However, since the variance is only a local approximation, we will add additional design restrictions on the

optimal processsor by adjoining constraints which control the ambiguity prone sidelobes of the GCC.

IV. The Hybrid Processor

The approach taken to optimize a combined envelope and nominal GCC processor is to penalize the candidate processor for poor local performance (small errors) while constraining it to supress its ambiguity prone sidelobes (large errors). The specific processor structure which we will investigate is displayed in Fig. 4.1, the GCC--GCC-envelope processor. Two envelope waveforms are obtained from the narrowband bandpass observations, x_1 and x_2. A common reference carrier, of frequency w_o (identical to the center frequency of the x_i's) and with uniform random phase, ϕ, is used to heterodyne the observations down to zero frequency. These envelopes are then mixed into the observations to form the composite bandpass and envelope waveforms, $\hat{x}_1$ and $\hat{x}_2$. The mixture of the envelopes into the observations is controlled by the parameter α where $\alpha \in [0,1]$. The resultant composite waveforms are subsequently processed by a classical GCC to form the peak estimate of time delay, $\hat{D}$. Note that the stationarity of the observations, and the random phase assumption on the heterodyne waveform, imply that the spectra, $\Phi_{11}(w)$ and $\Phi_{22}(w)$, and cross-spectrum, $\Phi_{12}(w)$, of $\hat{x}_1$ and $\hat{x}_2$ are simply

$$\Phi_{ij}(w) = \alpha^2 G_{ij}(w) + (1-\alpha)^2 G_{ij}(|w|-w_o) , \qquad i, j = 1, 2 \qquad (7)$$

Here, as in Chapter 2, G_{ij} denote the various auto- and cross-spectra of the original observations x_1 and x_2.

We will consider the following particular penalty function on the GCC weight $W(w)$ (local performance measure)

$$\rho(W) = var_L(\hat{D}) + \frac{1}{w_o^2} SNR_1^{-1} \qquad (8)$$

where var_L denotes the local variance, Eq. (23) of Chapter 3, and SNR_1 is the SNR introduced in Chapter 2, Eq. (12), and discussed in the previous section in connection with the BLB, Eq. (4). Note the above measures are with respect to the estimate yielded by the hybrid GCC-GCC-envelope-processor.

The specific combination of processor structure and penalty function , Fig. 4.1 and Eq. (8), was chosen because of our ability to emphasize locally optimal performance at either high or low SNR by choice of α. Indeed, for $\alpha = 1$, Eq. (8) takes its global minimum when the locally optimal HT weight, $W_{HT}(w)$ in Eq. (6), is used in place of $W(w)$ in Fig. 4.1, while for $\alpha = 0$, the minimum is attained for the HT implementation of the GCC envelope processor, i.e. $W(w) = W_{HT}(|w|-w_o)$. Thus α can be considered a parameter which adjusts the sensitivity of the penalty function to the SNR operating point of the hybrid processor.

In addition to local performance, characterized by $\rho(W)$, we are interested in good large error performance of the hybrid GCC with filter W. In this aim we impose the following inequality constraints on the envelope of the GCC output, R^g in Fig. 4.1.

$$\mathbf{E}[\frac{d}{d\tau} \ R^g(\tau)|_{\tau=D}] \le -K \tag{9}$$

$$\mathbf{E}[R^g(D+\frac{n2\pi}{w_o})] \le r_n \ , \quad n = -N, .\,. , 0, .\,. N \tag{10}$$

In (9), K is the magnitude of the curvature of the output of the hybrid processor at the true delay. Therefore (9) constrains the resolution of the peak at D relative to small errors. Eq. (10), on the other hand, constrains the resolution relative to large errors by maintaining low magnitude sidelobes (low magnitude envelope curvature). The particular form of constraints (9) and (10), for the GCC with weight $W(w)$, can be derived from Chapter 2, Eq. (4), and the assumptions contained therein, by identifying

$$E[\frac{d}{d\tau} \; R^{g}(\tau)|_{\tau=D}] = \frac{1}{2\pi} \int_{-\infty}^{\infty} w^{2}|G_{12}(w)| \, W(w) \, dw \qquad (11)$$

$$E[R^{g}(D+\frac{n2\pi}{w_{o}})] = \frac{1}{2\pi} \int_{-\infty}^{\infty} |G_{12}(w)| \, W(w) e^{jw\frac{n2\pi}{w_{o}}} \, dw \qquad (12)$$

The expressions for the GCC local variance and SNR_{1} are obtained from Eq. (23) of Chapter 3 and Eq. (12) of Chapter 2, by substituting the present input spectra, of Eqs. (7) and (9), in place of the spectra contained in formulas (23) and (12) of Chapter 3 and Chapter 2. The results are

$$var_{L}(\hat{D}) = \frac{\dfrac{1}{2\pi T} \displaystyle\int_{-\infty}^{\infty} w^{2} F(w) |W(w)|^{2} dw}{\left[\dfrac{1}{2\pi} \displaystyle\int_{-\infty}^{\infty} w^{2} G(w) W(w) dw\right]^{2}} \qquad (13)$$

and

$$SNR_{1}^{-1} = \frac{\dfrac{1}{2\pi T} \displaystyle\int_{-\infty}^{\infty} F(w) |W(w)|^{2} dw}{\left[\dfrac{1}{2\pi} \displaystyle\int_{-\infty}^{\infty} G(w) W(w) dw\right]^{2}} \qquad (14)$$

where the following spectral products have been defined in terms of the spectra of $\hat{x}_{1}$ and $\hat{x}_{2}$, $\Phi_{ij}(w)$, i,j = 1,2

$$G(w) = |\Phi_{12}(w)| \qquad (15)$$

and

$$F(w) = \Phi_{11}(w)\Phi_{22}(w) - G^{2}(w) \qquad (16)$$

In the present case we are interested in highly narrowband bandpass signal spectra. In this case the substantial portion of signal energy lies within the narrow frequency region $|w| \in [w_{o} -B/2, w_{o} +B/2]$. For the proposed hybrid implementation of Fig. 4.1, this implies that the inputs to the GCC , $\hat{x}_{1}$ and $\hat{x}_{2}$, will

have negligible signal related energy outside of narrow regions centered about w_o and zero frequencies. Certainly any reasonable weight $W(w)$ will attenuate the components of the GCC inputs outside of these regions. Consequently, as will be seen, the optimization of the bandpass and lowpass partions of $W(w)$ can be decoupled. For a typical $W(w)$ see Fig. 4.2.

Inequality constrained minimization problems can be attacked in different ways, see e.g. [6]. Since the object here is to quickly arrive at an optimal structure, we will assume that the so called *Kuhn–Tucker* conditions are satisfied for the constrained minimization of $\rho(W)$, (8) subject to (9) and (10). In this case a necessary condition for a given $W(w)$ to be optimal, in the sense of minimizing ρ over the constraint region defined by (9) and (10), is that it be a stationary point of the Euler-Lagrange form

$$(17)$$

$$var_L(\dot{D}) + \frac{1}{w_o^2} \; SNR_1^{-1} + \varsigma_o \, \mathbf{E}[\frac{d}{d\tau} \; R^g(\tau)|_{\tau=D}] + \sum_{n=-N}^{N} \lambda_i \, \mathbf{E}[R^g(D+\frac{n2\pi}{w_o})]$$

Where ς_o and λ_i $i=-N,..,0,..N$ are constant multipliers which are yet to be determined, and depend on the constraint regions. The form (17) can be written explicitly as an integral over frequency of the functions of observation spectra, Eqs. (15) and (16), and the GCC weight $W(w)$

$$\int_{-\infty}^{\infty} dw \; w^2 F(w)\frac{1}{K^2} \; |W(w)|^2 + F(w)\frac{1}{r_o^2} \; |W(w)|^2 \qquad (18)$$

$$+ \, \varsigma_o \, w^2 G(w) W(w) + \sum_{n=-N}^{N} \lambda_n \, e^{jw\frac{n2\pi}{w_o}} \; G(w) W(w)$$

Now express $F(w)$ and $G(w)$ in (18) in terms of the spectra $\Phi_{ij}(w)$ (See Eqs. (15), (16)). These quantities can in turn be expressed in terms of the disjoint spectra $G_{ij}(w)$ and $G_{ij}(|w|-w_o)$ (See Eq. (7)) to yield the equivalent form to (18)

$$\int_{-\infty}^{\infty} \alpha^2 \left[\frac{1}{K^2} \ (w^2 + \frac{K^2}{r_o^2}\ )F_o\left(w\right)|W\!\left(w\right)|^2 + (\varsigma_o\ w^2 + \sum_{n=-N}^{N} \lambda_n\ e^{jw\frac{n2\pi}{w_o}}\ )G_o\left(w\right)W\!\left(w\right) \right]$$

$$(19)$$

$$+ (1 - \alpha)^2 \left[\frac{1}{K^2} \ (w^2 + \frac{K^2}{r_o^2}\ )F_e\left(w\right)|W\!\left(w\right)|^2 + (\varsigma_o\ w^2 + \sum_{n=-N}^{N} \lambda_n\ e^{jw\frac{n2\pi}{w_o}}\ )G_e\left(w\right)W\!\left(w\right) \right]\ dw$$

where F_o and G_o are products of the nominal bandpass spectra of the observations

$$F_o\left(w\right) = G_{11}(w)G_{22}(w)(1 - |\gamma_{12}(w)|^2)\ , \qquad G_o\left(w\right) = |G_{12}(w)| \qquad (20)$$

and F_e and G_e are the lowpass equivalent spectra to the above

$$F_e\left(w\right) = F_o\left(|w| - w_o\right)\ , \qquad G_e\left(w\right) = G_o\left(|w| - w_o\right) \qquad (21)$$

It has been observed that $W\!\left(w\right)$ is significantly different from zero only on the support of G_o, F_o and G_e, F_e, which are highly concentrated in $|w| \in [\ w_o - B/2,\ w_o + B/2]$ and $w \in [-B/2,\ B/2]$ respectively. As a consequence, the specification of $W\!\left(w\right)$ over the bandpass frequency region can be made independently of that of $W\!\left(w\right)$ over the lowpass frequency region (the two terms in brackets, [..], in (19) are non-zero over essentially disjoint frequency regions). (19) can therefore be decomposed into two seperate forms to be minimized, the one involving the bandpass spectra, F_o, G_o and the other involving the lowpass F_e, G_e.

Formal differentiation of the resulting equations with respect to $W\!\left(w\right)$ yields the following necessary conditions for $W\!\left(w\right)$ to be an optimal GCC weight, which we denote $W_{opt}\left(w\right)$

$$(22)$$

$$W_{opt}\left(w\right) = \left[\frac{G_o\left(w\right)}{F_o\left(w\right)}\ + \frac{G_e\left(w\right)}{F_e\left(w\right)}\ \right] \frac{K^2}{w^2 + a^2}\ \left[\varsigma_o\ w^2 + \sum_{n=-N}^{N} \lambda_n\ e^{-jw\frac{n2\pi}{w_o}}\ \right]$$

Here $a^2 = K^2/r_o{}^2$. Observe that λ_o can be expressed as $\lambda_o = \lambda' + \varsigma_o\, a^2$, for some constant λ'. Hence, without loss of generality, we can rewrite Eq. (22) in its final form

$$W_{opt}(w) = \left[\frac{G_o(w)}{F_o(w)} + \frac{G_e(w)}{F_e(w)}\right]\left[K + \frac{1}{w^2 + a^2}\sum_{n=-N}^{N}\lambda_n\, e^{-jw\frac{n2\pi}{w_o}}\right] \quad (23)$$

where we have ignored any constant scaling of W_{opt}.

From the defining relations of G_o, F_o, G_e, and F_e, Eqs. (20) and (21), recognize the locally optimal HT weights $W_{HT}(w)$ and $W_{HT}(|w|-w_o)$ for the observations and their envelopes respectively

$$W_{HT}(w) = \frac{G_o(w)}{F_o(w)} \quad (24)$$

$$W_{HT}(|w|-w_o) = \frac{G_e(w)}{F_e(w)} \quad (25)$$

The sum $G_o/F_e + G_e/F_e$, in Eq. (23), will be denoted the locally optimal hybrid weighting $W_{lo}(w)$.

The structure of the optimal hybrid processor is displayed in Fig. 4.3 in block diagram format. Here the notation TDL denotes the implementation of the sum of weighted exponentials in (23) as a tapped delay line. The processor reduces to the following sequence of operations on the composite waveform $\hat{x}_1$ and $\hat{x}_2$. First locally optimal processing is performed on the composite waveforms using the Hannan-Thomson locally optimal (hybrid) weighting, W_{lo}; the sum of (24) and (25), in a conventional GCC. The output of the GCC is then passed through a two branch network. One of these branches, the lower one in Fig. 4.3, simply implements the locally optimal estimator by passing a scaled version of the output of W_{lo} directly to the output, R^g. The other branch is more interesting. It takes the ambiguity prone output of W_{lo} and performs

equalization, via a tapped delay line structure, to minimize sidelobe activity. The resultant waveform is then lowpass filtered to remove spurious high frequency peaks, which could cause ambiguity. Finally the equalized and filtered waveform is combined with the locally optimal GCC output to form the peak generating sample cross-correlation function, R^g.

As the above development suggests, the final configuration is intuitively satisfying. The next step would be to simulate the processor and compare it to others which have been designed for narrowband lowpass signals, e.g. the Hassab-Boucher (HB), reference [2] of Chapter 2, and the simple envelope GCC suggested in [4]. This will be a subject of a future study.

V. Conversion to a Pseudo-linear Form.

Here some ideas and preliminary results are presented concerning a procedure which transforms the multi-waveform, non-linear time delay estimation problem: estimate $D \in [-D_m, D_m]$ from

$$x_1(t) = s(t) + n_1(t) \tag{26}$$

$$t \in [0, T]$$

$$x_2(t) = s(t-D) + n_2(t)$$

into the single waveform, pseudo-linear, time delay estimation problem

$$H(t) = D + \eta(t) \qquad t \in [t_1, t_2] \tag{27}$$

where, under certain assumptions on the signal, $s(t)$, $\eta(t)$ is a non-stationary noise process, t_2 is a constant, and t_1 is a parameter determined by a "genie" algorithm, to be specified in a moment. We use the "pseudo-linear" nomenclature for (27) since t_1 depends on D in practice. The transformation yields a sufficient statistic, so that the resultant statistic (related to $H(t)$ above) contains as much information on the delay parameter as the original observations (26).

Although the technique can be applied to the original Gaussian signal case (In which case D gets multiplied by a random process, dependent on D, in (27)), the treatment to follow will be restricted to the deterministic signal case. However D may be a random variable. Two specific signal characteristics will be considered: 1) $s(t)$ of finite time duration (finite energy), T_e ; and 2) $s(t)$ periodic with period T_S and low duty cycle. For 1), the signal, and its delayed version, are assumed to be entirely contained in the observation interval $[0, T]$. For 2), the delay is restricted to the interval $[-D_m, D_m] = [-T_S/2, T_S/2]$.

As before, form the sample cross-correlation function between x_1 and x_2

$$\hat{R}_{12}(\tau) = \frac{1}{T} \int_0^T x_1(t)x_2(t+\tau)dt \quad , \qquad \tau \in [-T, T] \tag{28}$$

Recall the decomposition of $\hat{R}_{12}$ in Chapter 2 , Eq. (5). The identical result holds here once the signal auto-correlation function is defined for the deterministic case

$$R_{ss}(\tau) = \int_0^T s(t)s(t+\tau)dt \tag{29}$$

for the finite energy signal, and

$$R_{ss}(\tau) = \frac{1}{T} \int_0^T s(t)s(t+\tau)dt \tag{30}$$

for the periodic signal. In defining (29) and (30) we have used the assumptions on the time extents and period of $s(t)$ and $s(t-D)$, discussed in the preceeding paragraph, to ensure that (29) and (30) are well defined (independent of T). The decomposition of (28) can then be expressed in the form

$$\hat{R}_{12}(\tau) = r(\tau-D) + R_n(\tau) \tag{31}$$

where R_n is the sample cross-correlation between the different noise and signal times noise terms in $\hat{R}_{12}$

$$R_n(\tau) = \frac{1}{T} \int_0^T \left(s(t)n_2(t+\tau) + n_1(t)s(t-D+\tau) + n_1(t)n_2(t+\tau) \right) dt \qquad (32)$$

and

$$\dot{r}(\tau) = \frac{1}{T} \; R_{ss}(\tau) \qquad (33)$$

for the finite energy case, and

$$\dot{r}(\tau) = R_{ss}(\tau) \qquad (34)$$

for the periodic signal.

It is straightforward to show that R_{ss}, in (29) and (30), takes its global maximum at D and is also symmetric about D. Therefore $r(\tau)$ can be regarded as a true autocorrelation function. This is in contrast to the random signal case, where in the decomposition of Chapter 2, Eq. (5), $\hat{R}_{ss}$ is only an asymptotic approximation to an auto-correlation function.

It is immediately verified that the cross-correlation noise, R_n in (32) is zero mean, as in the random signal case (See Chapter 3). A bit more tedious, is the calculation of the covariance, $\sigma(t,\tau)$, of R_n, which is, however, derived analogously to the derivation in Chapter 3, Appendix C, for the random case. The result is

$$(35)$$

$$\sigma(t,\tau) = \frac{1}{T} \int_{-T}^{T} \left[R_{n1}(\xi)R_{n2}(\xi+t-\tau)(1 - \frac{|\xi|}{T}) + R_{n2}(\xi+t-\tau)r^{(1)}(\xi) + R_{n1}(\xi)r^{(2)}(\xi,t,\tau,D) \right] d\xi$$

where

$$r^{(1)}(\xi) = \frac{1}{T} \int_0^{T-|\xi|} s(\nu)s(\nu+|\xi|)d\nu \qquad (36)$$

and

$$r^{(2)}(\xi,t,\tau,D) = \frac{1}{T} \int_0^{T-|\xi|} s(\nu+t-D+|\xi|)s(\nu+t-D+\tau)d\nu \qquad (37)$$

and R_{ni} , i $=$ 1,2 , are the auto-correlation functions of the observation noise, n_i in (26). Note, as in the random signal case, the covariance is a function of D and depends on t and τ in addition to $t - \tau$.

Define the constant L

$$L = D_m + T_c \tag{38}$$

and the (linear) function of $\hat{R}_{12}$

$$H(t) = \int_{-L}^{t} \tau \hat{R}_{12}(\tau) d\tau \ , \qquad t \in [-L,L] \tag{39}$$

Note that H is invertible ($\hat{R}_{12}(\tau) = \frac{1}{\tau} \ dH(\tau)/d\tau$). Use the decomposition, (31), in (39) to obtain

$$H(t) = \int_{-L}^{t} \tau r(\tau-D) d\tau + \int_{-L}^{t} \tau R_n(\tau) d\tau \ , \qquad t \in [-L,L] \tag{40}$$

Now make a change of variable in the first integral on the right of the equality (40) and identify the noise term η

$$H(t) = D \int_{-L-D}^{t-D} r(\tau) d\tau + \ + \int_{-L-D}^{t-D} \tau r(\tau) d\tau + \eta(t) \tag{41}$$

where

$$\eta(t) = \int_{-L}^{t} \tau R_n(\tau) d\tau \tag{42}$$

The covariance of η can be calculated straightforwardly from Eq. (35)

$$\sigma'(t,\tau) = \int_{-L}^{t} d\nu \int_{-L}^{\tau} d\mu \ \sigma(\nu,\mu) \tag{43}$$

In its present form, Eq. (41) is not quite linear in D. However, if $r(\tau)$ is of width T_c , then, because of the symmetry of $r(\tau)$, the first two integrals in (41) are zero for $-L \leq t \leq D-T_c$, and the first integral is a known constant, the

second zero, for $D+T_c \leq t \leq L$. As a consequence, neglecting the influence of the D-dependent noise, R_n, $H(t)$ is linear in D over the latter interval, and independent of D over the former.

For simplicity assume that $r(\tau)$ is a narrow pulse (broadband signal case). Then make the approximation

$$H(t) = Du(t-D) + \eta(t) \ , \qquad t \in [-L,L] \qquad (44)$$

where $u(\tau)$ is the unit step function, taking values zero, for negative τ, and one, for positive τ. For unknown, deterministic D, the maximum-likelihood estimate of D can be specified under a Gaussian assumption on the zero mean noise, $\eta(t)$. The resultant (classical) estimator is illustrated in Fig. 5.1, which seeks to minimize the (prewhitened) residuals between $H(t)$ and $\hat{D}u(t-\hat{D})$. Observe that the estimate $\hat{D}$ updates the "inverse filter": $R_n^{-1}(t,\tau)$. Here R_n^{-1} is the inverse kernel associated with the noise covariance of (43), defined such that

$$\delta(\tau-t) = \int_{-L}^{L} R_n^{-1}(\tau,\nu)\sigma'\ (\nu,t)d\nu \qquad (45)$$

For a low SNR condition and white noises of power $N_o/2$, the estimator of Fig. 5.1 no longer needs to update the inverse kernel R_n^{-1}. This comes from the D-independent form of the covariance, Eq. (43), of η for low SNR

$$\sigma'\ (\tau,t) = N_o{}^2\frac{[min(\tau,t)]^3 + L^3}{6T} \qquad (46)$$

It remains to be determined if the time varying inverse filter, R_n^{-1} of (45), is of simple enough form to be a practical implementation.

The maximum likelihood estimator (MLE) is an optimal implementation. However, the "signal", D, is only present for a portion of the time interval $[-L,L]$. Consequently the MLE, implemented as above, over-utilizes the data, part of which may be purely observation noise. An alternate, possibly sub-optimal, pro-

cedure would involve the implementation of a causal, joint estimation-detection scheme to estimate D from $H(t)$. Essentially one would begin observing $H(t)$ at $t = L$, and form an estimate of D, $\hat{D}(t)$, as t decreases to $t = L-T_c$ (In this interval it is known that the signal D is present in the observations (41)). For $t \leq L-T_c$, the estimate $\hat{D}(t)$ could be used as evolving *a priori* information in a hypothesis test of the presence or absence of the "signal" D in additive noise $\eta(t)$. When the detection algorithm indicates the absence of D in the observations, $H(t)$, the estimation procedure is terminated. The above approach converts the estimation problem, (41), into the linear problem of (27), where the "genie" is identified with the detection algorithm, and $t_2 = L$. The details of the proposed estimation-detection scheme remain to be determined by future study.

Some insight into the performance of the MLE of Fig. 5.1, for low SNR and white noise, may be obtained from considering the performance of the following simple estimate of D

$$\hat{D} = \frac{H(L)}{\int\limits_{-L}^{L} r(\tau)d\tau} \tag{47}$$

where L is defined in (38). Note, it must be assumed that the integral in (47) is greater than zero for the estimate to be well defined. This is true if $r(\tau)$ is strictly positive definite, or if there is a DC component in the signal auto-correlation function. Since only a single point of $H(t)$, $H(L)$, is used in (47) and no inverse filtering is implemented, $\hat{D}$ of (47) is sub-optimal, in comparison to the MLE. The estimate (47) is unbiased since the mean of η, Eq. (42), is zero and in (41)

$$\int\limits_{-L-D}^{L-D} r(\tau)d\tau = \int\limits_{-L}^{L} r(\tau)d\tau \quad \text{and} \quad \int\limits_{-L-D}^{L-D} \tau r(\tau)d\tau = \int\limits_{-L}^{L} \tau r(\tau)d\tau = 0 \tag{48}$$

by virtue of the definition of L, given in (38), and the fact that the symmetric function $r(\tau)$ is zero outside of $\tau \in [-T_c, T_c]$.

As a simple example for the finite energy case, consider the case of a triangular shaped signal auto-correlation of width $2T_c$ and maximum S at zero. In this case

$$\int_{-L}^{L} r(\tau)d\tau = \frac{ST_c}{T} \tag{49}$$

and the low SNR variance of the estimate (47) is obtained immediately from (41), (43) and (46)

$$var[\hat{D}] = \frac{\sigma'(L,L)}{\left[\int_{-L}^{L} r(\tau)d\tau\right]^2} = \frac{LT(\frac{L}{T_c})^2}{6SNR^2} \tag{50}$$

Recall, $L = D_m + T_c$ and $SNR = S/N_o$. Eq. (50) indicates some of the important performance factors. There are tradeoffs between D_m, T and T_c. In principle, unlike the random signal case, performance does not decrease with increasing T since the noise dominates over all but a time interval of length T_c. For the same reason, in contrast to the behavior of the GCC at high SNR for random signals, the broader the signal auto-correlation, over the *a priori* region (small L/T_c), the lower the variance. Increasing D_m, on the other hand, has analogous effects on performance as in the random signal case; it introduces more uncertainty into the delay parameter over time, thus increasing the error.

For the periodic signal case, let the auto-correlation function consist of copies of the triangular auto-correlation function studied above, with interspacing T_S, where $T_c \ll T_S$. With the assumption $T_S > D_m + T_c$ (no ambiguity because of phase), the variance for low SNR becomes

$$var[\hat{D}] = \frac{\frac{L}{T}(\frac{L}{T_c})^2}{6SNR^2} \tag{51}$$

The form of (51) is identical to (50) except for the dependence on T. Here the variance decreases with increasing T since infinite signal observation time is available.

For the deterministic (radar) signal, the expressions for the variance of the conventional time delay estimator known to us are valid only for small errors, i.e. high SNR, or for special signal situations (e.g. white or Markovian signals and noises [7]). Since (51) and (50) are valid only for low SNR and involve additional parameters absent in the conventional estimator (e.g. L), comparison of the performance of the linearized procedure, discussed here, with that of conventional techniques is difficult. Further research into the relative merits of this estimation scheme is necessary, before any general conclusions can be drawn.

VI. Conclusions

The two signal processing schemes introduced in this chapter take interesting forms. However, additional study is needed to quantify the relative performance of these methods with respect to the conventional techniques.

For the hybrid GCC-GCC envelope processor, a deeper analysis into the necessary conditions for the optimality of the structure of Fig. 4.3 for the constrained minimization is necessary. In particular, one needs to know how to choose the various design parameters such that a). a solution of the constrained optimization exists, and if so b). the implementation of the resultant processor is sufficiently simple. For example, the choice of α, the observation versus envelope mixture parameter, should correspond to the (perhaps changing) priorities of high vs. low SNR performance. Furthermore, the number and magnitude of the constraint boundaries will certainly have an impact on the performance, complexity and even the existence of the optimal estimator.

The linearization method, discussed in Section V, is only (approximately)

linear for certain characteristic (deterministic) signal forms. For the general case the transformed cross-correlation function remains non-linear, a sum of a linear function of D and the product of two linear functions of D. However, the non-linearity may be a simple non-linearity (i.e. easy to invert). This will be signal dependent, and suggests the investigation of signal design for this technique. The performance of the method, in any case, remains to be thoroughly investigated.

REFERENCES

Chapter 6

1. C.H. Knapp and G.C. Carter, "The generalized correlation method for estimation of time delay," *IEEE Trans. on ASSP*, Vol. ASSP-24, No. 4, pp. 320-327, 1976.

2. B. Friedlander, "On the Cramer-Rao bound for time delay and doppler estimation," *IEEE Trans. on Inform. Theory*, Vol. IT-30, No. 3, pp. 575-580, 1984.

3. A.J. Weiss and E. Weinstein, "Fundamental limitations in passive time delay estimation - Part I: narrowband systems," *IEEE Trans. on ASSP*, Vol. ASSP-31, No. 2, pp. 472-485, 1983.

4. S.K. Chow and P.M. Schultheiss, "Delay estimation using narrowband processes". *IEEE Trans. on ASSP*, Vol. ASSP-29, No. 3, pp. 478-484, 1981.

5. E.W. Barankin, "Locally best unbiased estimates," *Annals of Math. Stat.*, Vol. 20, pp. 477-501, 1949.

6. D.G. Luenberger, *Optimization by Vector Space Methods*, Wiley, N.Y., 1969.

7. V. DiFranco, *Radar Detection* Prentice Hall, Englewood-Cliffs, N.J., 1968

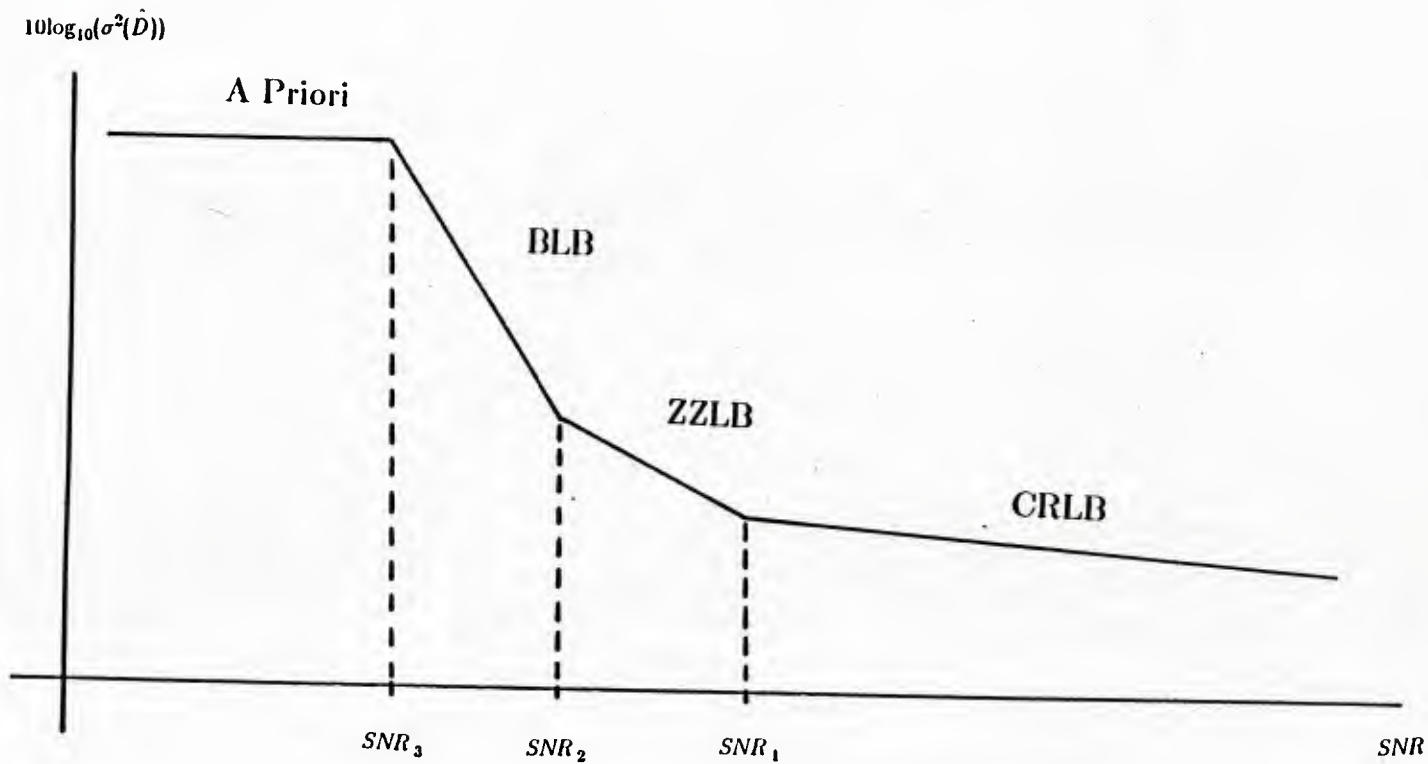Composite Lower Bound on Variance
for Narrowband Signal Spectrum

Fig. 2.1

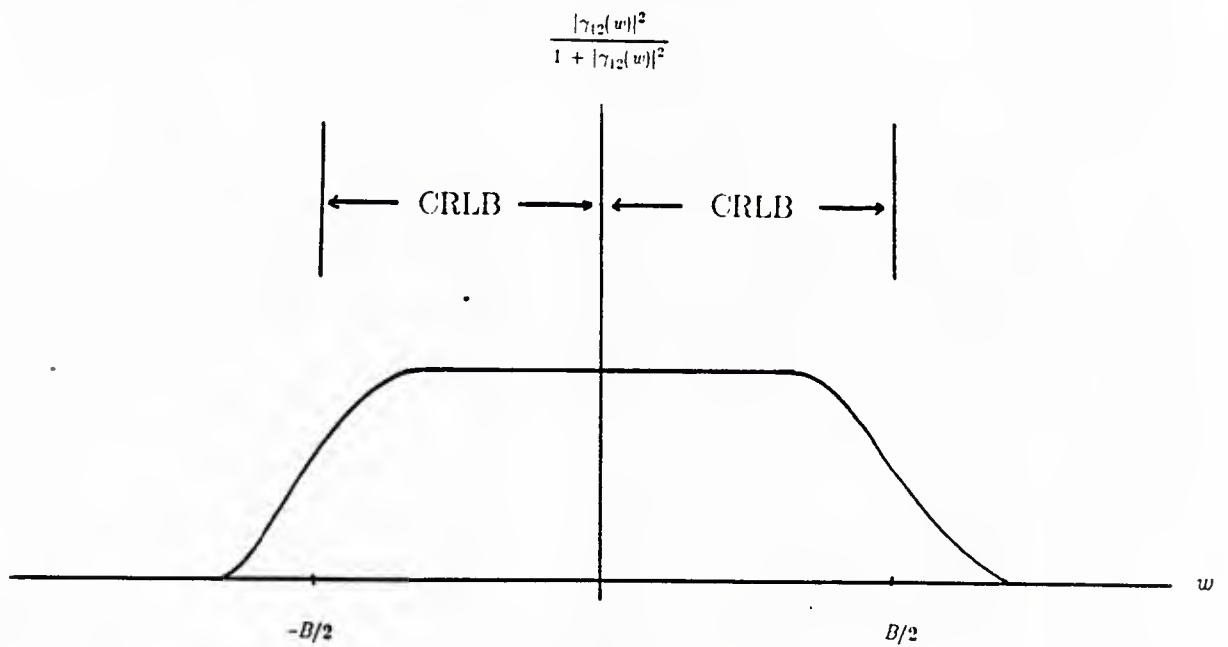## CRLB Measures Frequency Occupancy of Lowpass Signal Spectrum

$$\frac{|\gamma_{12}(w)|^2}{1 + |\gamma_{12}(w)|^2}$$

$\leftarrow$ CRLB $\rightarrow$ | $\leftarrow$ CRLB $\rightarrow$

$-B/2$ $B/2$ w

Fig. 2.2a

BLB Measures Frequency Occupancy of Bandpass Signal Spectrum

$$\frac{|\gamma_{12}(w)|^2}{1 + |\gamma_{12}(w)|^2}$$

$\leftarrow$ CRLB $\rightarrow$ | $\leftarrow$ CRLB $\rightarrow$

$\leftarrow$ BLB $\rightarrow$ $\leftarrow$ BLB $\rightarrow$
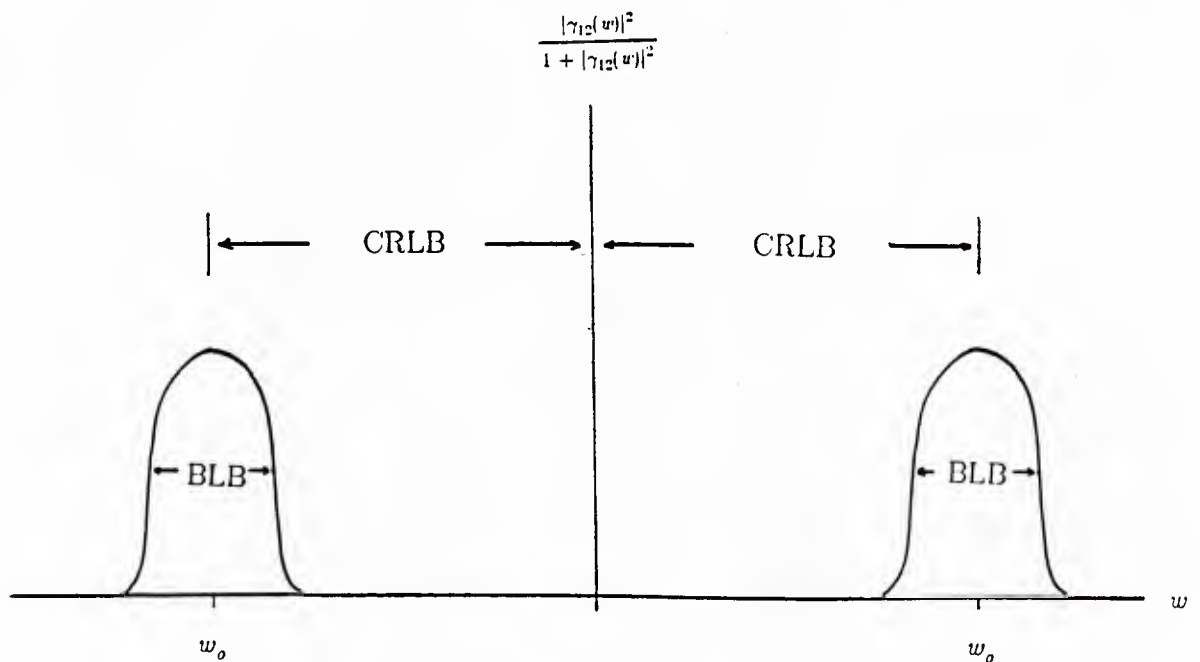
w_o w_o w

Fig. 2.2b

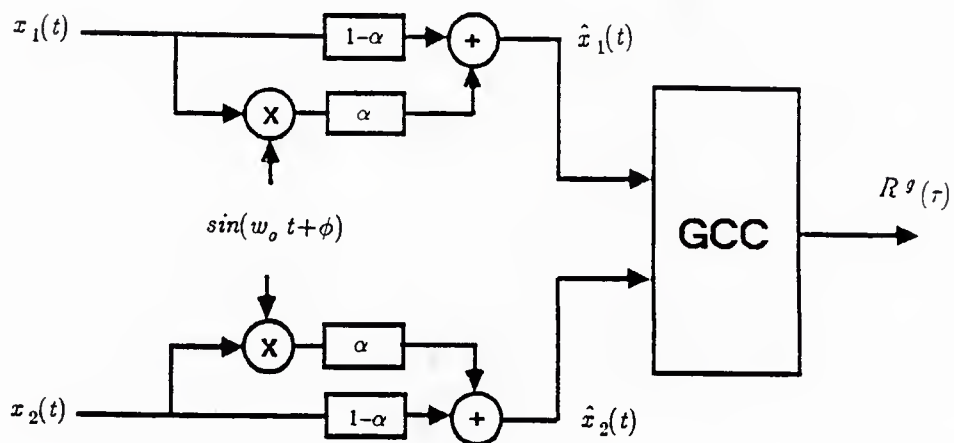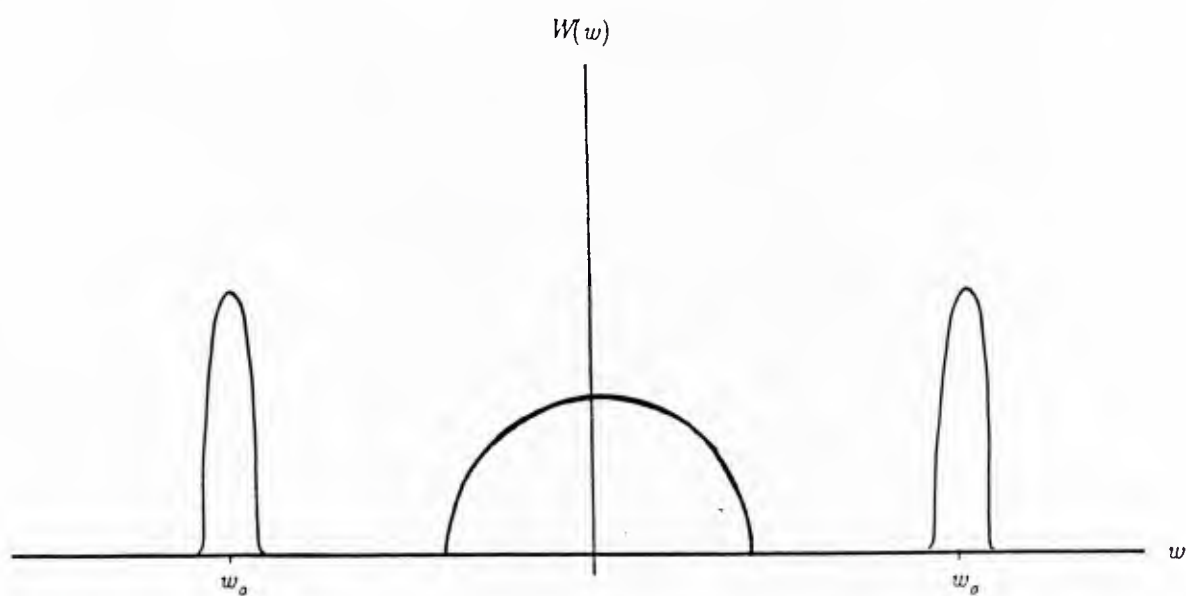## Hybrid GCC-GCC Envelope Processor

Fig. 4.1

Typical GCC weight $W(w)$ for Hybrid GCC

Fig. 4.2

202

Optimal Hybrid GCC-GCC Envelope Processor
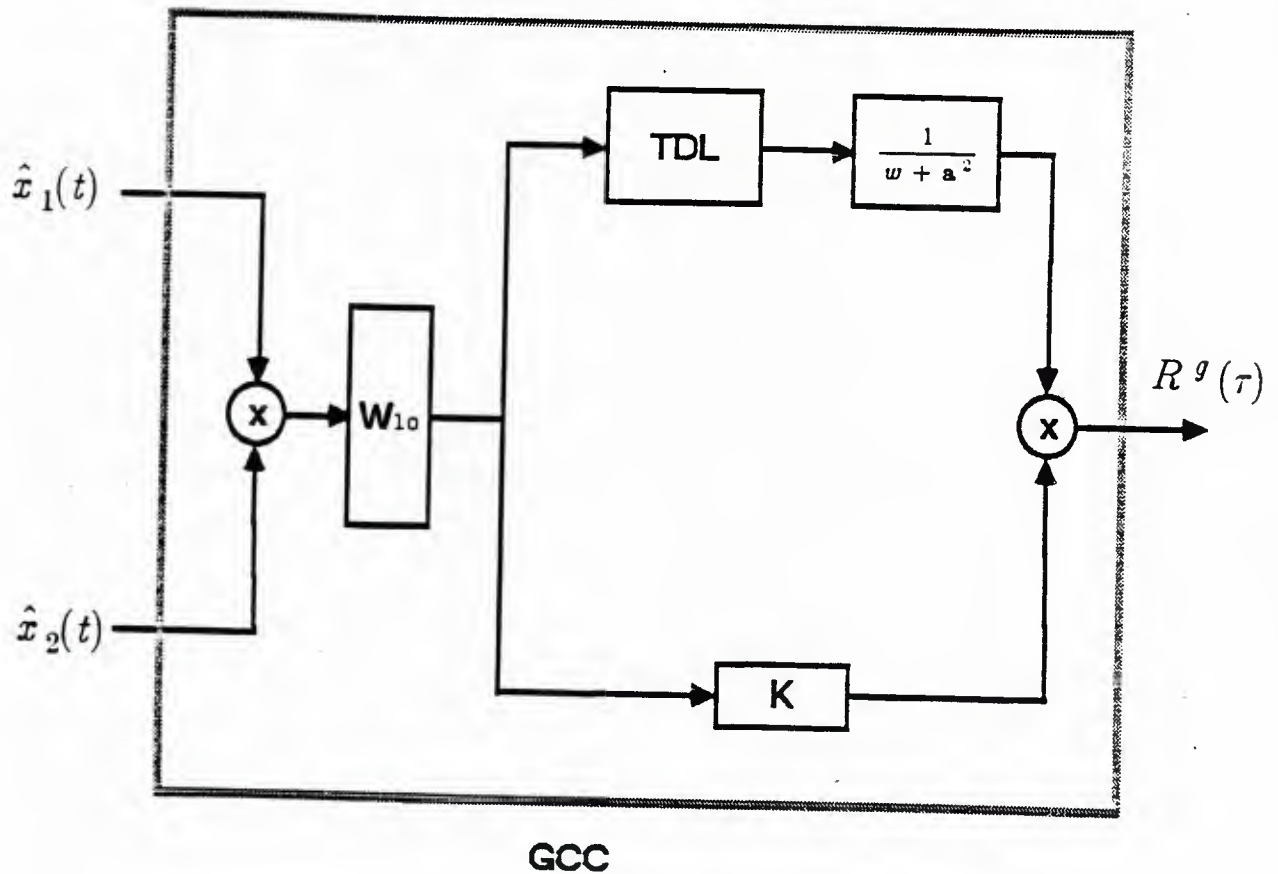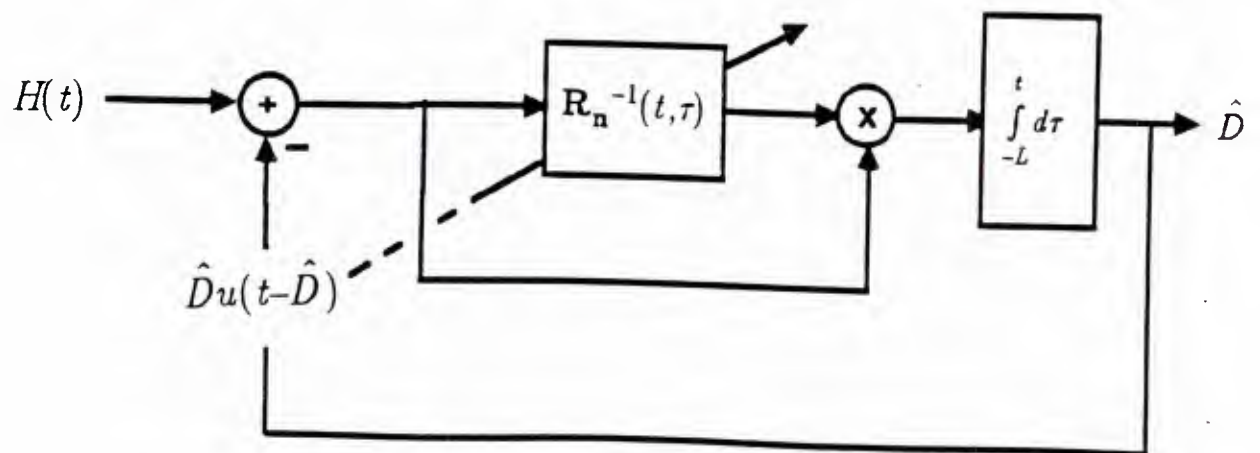

Fig. 4.3

Maximum Likelihood Delay Estimator given $H(t)$

Fig. 5.1

Chapter 7
Conclusion

Two sensor, passive time delay estimation has been considered in this dissertation. Both the development of new estimators and the characterization of estimator performance have been of interest. Of prime concern, however, have been two situations for which the conventional correlation type delay estimators display poor performance: unknown observation spectra, and low time-bandwidth product (BT).

In Chapter 1, cross-correlation of the least mean square estimates (LMSE) of the signal and its delayed version gave a new correlation type delay estimate, the Wiener Processor (WP). For partially unknown observation spectra, replacement of the LMSEs by robust (min-max) estimates in the WP gave the Robust Wiener Processor (RWP). The RWP represents a first step in obtaining min-max time delay estimates for the difficult case of unknown spectra.

Chapters 3 and 6 introduced alternate processing strategies designed to improve upon the conventional techniques for narrowband signals (low BT). The Center of Symmetry Estimate (CSE) was proposed in Chapter 3 as a way to desensitize the estimate to large errors (outliers). In Chapter 6, a combined signal and signal envelope correlation type processor was optimized to yield a novel estimator structure for narrowband bandpass signals.

In Chapters 4 and 5 the attention was focused on the theory of upcrossing probabilities and their application to modeling the large error performance of correlation type time delay estimates. Chapter 4 represents an extension of the existing results on the asymptotic distribution of the level crossings of a stationary, continuous time random process to the non-stationary case. Then, a level crossing model for large delay estimator errors was taken in Chapter 5. The

approach taken in Chapter 5 is more general than the previous large error approximations in that it can be applied to the case of bandpass as well as lowpass signals. Motivated by the results of Chapter 4, approximations to the probability of large error and the estimator variance were derived. For large BT, and lowpass signal spectra, the approximations are in good agreement with the simulated performance of the cross-correlator. For bandpass spectra they are consistent with the relevant performance predictions of the Ziv-Zakai lower bound on achievable error.

At the conclusion of this research several avenues are open for future study. Extensions to more general observation models than that treated here warrant investigation. Of special interest is the case of correlated noise and signal components, which occurs in multipath and reverberation environments. Also of interest is the multi-sensor situation. In both of the above scenarios the pairwise sensor cross-correlations cease to be sufficient statistics for the inter-sensor delays. Therefore a substantially different approach than that taken in this thesis may be necessary. Likewise, the case of a time evolving delay parameter, as occurs in tracking, introduces further complexity into the estimation problem. However, in the latter case the sample cross-correlation function remains sufficient for the delay. Consequently, in this case, some extensions of the present work may be possible.

OFFICE OF NAVAL RESEARCH
STATISTICS AND PROBABILITY PROGRAM

BASIC DISTRIBUTION LIST
FOR
UNCLASSIFIED TECHNICAL REPORTS

FEBRUARY   1982

	Copies		Copies
Statistics and Probability Program (Code 411(SP)) Office of Naval Research Arlington, VA 22217	3	Navy Library National Space Technology Laboratory Attn: Navy Librarian Bay St. Louis, MS 39522	1
Defense Technical Information Center Cameron Station Alexandria, VA 22314	12	U. S. Army Research Office P.O. Box 12211 Attn: Dr. J. Chandra Research Triangle Park, NC 27706	1
Commanding Officer Office of Naval Research Eastern/Central Regional Office Attn: Director for Science Barnes Building 495 Summer Street Boston, MA 02210	1	Director National Security Agency Attn: R51, Dr. Maar Fort Meade, MD 20755	1
Commanding Officer Office of Naval Research Western Regional Office Attn: Dr. Richard Lau 1030 East Green Street Pasadena, CA 91101	1	ATAA-SL, Library U.S. Army TRADOC Systems Analysis Activity Department of the Army White Sands Missile Range, NM 88002	1
U. S. ONR Liaison Office - Far East Attn: Scientific Director APO San Francisco 96503	1	ARI Field Unit-USAREUR Attn: Library c/o ODCSPER HQ USAEREUR & 7th Army APO New York 09403	1
Applied Mathematics Laboratory David Taylor Naval Ship Research and Development Center Attn: Mr. G. H. Gleissner Bethesda, Maryland 20084	1	Library, Code 1424 Naval Postgraduate School Monterey, CA 93940	1
Commandant of the Marine Coprs (Code AX) Attn: Dr. A. L. Slafkosky Scientific Advisor Washington, DC 20380	1	Technical Information Division Naval Research Laboratory Washington, DC 20375	1
		OASD (I&L), Pentagon Attn: Mr. Charles S. Smith Washington, DC 20301	1

Director
AMSAA
Attn: DRXSY-MP, H. Cohen
Aberdeen Proving Ground, MD   1
  21005

Dr. Gerhard Heiche
Naval Air Systems Command
    (NAIR 03)
Jefferson Plaza No. 1
Arlington, VA   20360   1

Dr. Barbara Bailar
Associate Director, Statistical
  Standards
Bureau of Census
Washington, DC 20233   1

Leon Slavin
Naval Sea Systems Command
    (NSEA 05H)
Crystal Mall #4, Rm. 129
Washington, DC   20036   1

B. E. Clark
RR #2, Box 647-B
Graham, NC   27253   1

Naval Underwater Systems Center
Attn: Dr. Derrill J. Bordelon
        Code 601
Newport, Rhode Island   02840   1

Naval Coastal Systems Center
Code 741
Attn: Mr. C. M. Bennett
Panama City, FL 32401   1

Naval Electronic Systems Command
    (NELEX 612)
Attn: John Schuster
National Center No. 1
Arlington, VA   20360   1

Defense Logistics Studies
  Information Exchange
Army Logistics Management Center
Attn: Mr. J. Dowling
Fort Lee, VA   23801   1

Reliability Analysis Center (RAC)
RADC/RBRAC
Attn: I. L. Krulac
      Data Coordinator/
      Government Programs
Griffiss AFB, New York 13441   1

Technical Library
Naval Ordnance Station
Indian Head, MD 20640   1

Library
Naval Ocean Systems Center
San Diego, CA 92152   1

Technical Library
Bureau of Naval Personnel
Department of the Navy
Washington, DC 20370   1

Mr. Dan Leonard
Code 8105
Naval Ocean Systems Center
San Diego, CA 92152   1

Dr. Alan F. Petty
Code 7930
Naval Research Laboratory
Washington, DC   20375   1

Dr. M. J. Fischer
Defense Communications Agency
Defense Communications Engineering
  Center
1860 Wiehle Avenue
Reston, VA 22090   1

Mr. Jim Gates
Code 9211
Fleet Material Support Office
U. S. Navy Supply Center
Mechanicsburg, PA 17055   1

Mr. Ted Tupper
Code M-311C
Military Sealift Command
Department of the Navy
Washington, DC 20390   1

Mr. F. R. Del Priori
Code 224
Operational Test and Evaluation
 Force (OPTEVFOR)
Norfolk, VA 23511   1